



Correction technique for Raman water vapor lidar signal-dependent bias and suitability for water vapor trend monitoring in the upper troposphere

D. N. Whiteman¹, M. Cadirola², D. Venable³, M. Calhoun³, L. Miloshevich⁴, K. Vermeesch⁵, L. Twigg⁵, A. Dirisu⁶, D. Hurst⁷, E. Hall⁷, A. Jordan⁷, and H. Vömel⁸

¹NASA/GSFC, Greenbelt, MD 20771, USA

²Ecotronics, LLC, Clarksburg, MD 20871, USA

³Howard University, Washington, DC 20059, USA

⁴Milo Scientific, LLC, Lafayette, CO 80026, USA

⁵SSTI, Lanham, MD 20706, USA

⁶Oak Ridge Associated Universities, Oak Ridge, 37381, TN, USA

⁷Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado 80309 and NOAA Earth System Research Laboratory, Global Monitoring Division, Boulder, 80305, Colorado, USA

⁸Lindenberg Observatory, Lindenberg, Germany

Correspondence to: D. N. Whiteman (david.n.whiteman@nasa.gov)

Received: 30 June 2011 – Published in Atmos. Meas. Tech. Discuss.: 12 December 2011

Revised: 22 October 2012 – Accepted: 1 November 2012 – Published: 28 November 2012

Abstract. The MOHAVE-2009 campaign brought together diverse instrumentation for measuring atmospheric water vapor. We report on the participation of the ALVICE (Atmospheric Laboratory for Validation, Interagency Collaboration and Education) mobile laboratory in the MOHAVE-2009 campaign. In appendices we also report on the performance of the corrected Vaisala RS92 radiosonde measurements during the campaign, on a new radiosonde based calibration algorithm that reduces the influence of atmospheric variability on the derived calibration constant, and on other results of the ALVICE deployment. The MOHAVE-2009 campaign permitted the Raman lidar systems participating to discover and address measurement biases in the upper troposphere and lower stratosphere. The ALVICE lidar system was found to possess a wet bias which was attributed to fluorescence of insect material that was deposited on the telescope early in the mission. Other sources of wet biases are discussed and data from other Raman lidar systems are investigated, revealing that wet biases in upper tropospheric (UT) and lower stratospheric (LS) water vapor measurements appear to be quite common in Raman lidar systems. Lower stratospheric climatology of water vapor is investigated both as a means to check

for the existence of these wet biases in Raman lidar data and as a source of correction for the bias. A correction technique is derived and applied to the ALVICE lidar water vapor profiles. Good agreement is found between corrected ALVICE lidar measurements and those of RS92, frost point hygrometer and total column water. The correction is offered as a general method to both quality control Raman water vapor lidar data and to correct those data that have signal-dependent bias. The influence of the correction is shown to be small at regions in the upper troposphere where recent work indicates detection of trends in atmospheric water vapor may be most robust. The correction shown here holds promise for permitting useful upper tropospheric water vapor profiles to be consistently measured by Raman lidar within NDACC (Network for the Detection of Atmospheric Composition Change) and elsewhere, despite the prevalence of instrumental and atmospheric effects that can contaminate the very low signal to noise measurements in the UT.

1 Introduction

Water vapor is one of the most important components of the atmosphere when considering atmospheric chemistry, radiation, dynamics and clouds. Water vapor profiles with sufficient accuracy and temporal and vertical resolution are needed to resolve variability in this important parameter in support of studies in all these areas. Climate model predictions indicate that water vapor will increase in the atmosphere as temperatures increase due to climate change (Held and Soden, 2006; Oman et al., 2008; Boers and Meijgaard, 2009; Whiteman et al., 2011b). The largest percentage changes are expected in the upper troposphere where model predictions show annual increases of up to 1% and more during the current century (Soden et al., 2005; Boers and Meijgaard, 2009; Whiteman et al., 2011b). For these reasons and others, significant effort has been put into both measurements and modeling of upper tropospheric water vapor to assess long-term trends in water vapor concentrations and thus address the consequences of future changes in atmospheric water vapor amounts. International efforts such as the Raman water vapor lidar initiative within the Network for the Detection of Atmospheric Composition Change (NDACC), www.ndsc.ncep.noaa.gov, and the Global Climate Observing System Reference Upper Air Network (GRUAN), www.gruan.org, are being established in response to the need to monitor water vapor trends in the atmosphere.

MOHAVE-2009 (Leblanc et al., 2011) was held in October 2009 at Table Mountain, CA, with the goal of characterizing a large suite of water vapor instrumentation used within NDACC. Information about the campaign can be found at <http://tmf-lidar.jpl.nasa.gov/campaigns/mohave2009.htm>. Here we report on the participation of the mobile system referred to as ALVICE (Atmospheric Laboratory for Validation, Interagency Collaboration and Education) in the MOHAVE campaign. A wet bias was found in the ALVICE Raman lidar upper tropospheric and lower stratospheric (UTLS) water vapor mixing ratio measurements. A correction that is based on lower stratospheric water vapor climatology is developed to address this wet bias and is applied to the ALVICE Raman lidar data processed for MOHAVE-2009. Reference to lower stratospheric climatology is found to be useful both for data quality control and for correction of Raman water vapor lidar measurements intended for scientific studies in the upper troposphere. The main body of this paper is devoted to the rationale for, and development of, the wet bias correction. Several appendices are also provided that give more of the traditional calibration/validation results during the MOHAVE-2009 campaign. These appendices include information on the performance of the corrected Vaisala RS92 radiosonde during MOHAVE-2009; the development of a new radiosonde calibration technique that reduces the influence of atmospheric variability on the derived lidar calibration constant; the quantification of the full uncertainty budget of Raman lidar water vapor mixing ratio measurements; the

development of specific data products for optimum comparison with radiosondes, satellites or other lidar systems; and a comparison of Integrated Precipitable Water (IPW) from lidar and GPS.

2 ALVICE

ALVICE consists of several measurement systems contained in a mobile environmentally-controlled trailer for deploying to field locations. The instrumentation now housed within the trailer includes an upgraded version of the Raman Airborne Spectroscopic Lidar (RASL) (Whiteman et al., 2007) mounted in an upward looking configuration for performing vertical profiling of water vapor, aerosols, temperature and clouds (Whiteman et al., 2010). The additional instrumentation includes a ventilated chamber, referred to as THPref (Temperature-Humidity-Pressure reference), which provides continuous, NIST-traceable surface data and permits characterizing radiosonde accuracy prior to launch, a SuomiNet (Ware et al., 2000) GPS instrument for integrated precipitable water (IPW) measurements, equipment for launching Vaisala RS92 and Internet radiosondes and cryogenic frost point hygrometers (CFH), and a stabilized calibration lamp that supports an independent calibration of a Raman water vapor lidar (Venable et al., 2011).

2.1 ALVICE lidar

The Raman Airborne Spectroscopic Lidar (RASL) mounted in an upward-looking configuration forms the heart of the lidar system contained within ALVICE and has been described previously in the literature (Whiteman et al., 2007, 2010). Since the time of those publications and prior to the MOHAVE-2009 campaign, the only significant lidar system modifications were the addition of rotational Raman temperature measurement capability and the use of a thermoelectrically cooled photomultiplier tube (PMT) for the water vapor channel measurements. Therefore, just a brief summary of the lidar instrumentation will be given here.

The Raman lidar contained in the ALVICE trailer uses an injection-seeded Continuum 9050 Nd:YAG laser operating at the frequency-tripled wavelength of 354.7 nm that emits pulses with approximately 300–350 mJ of energy at a rate of 50 Hz. The backscattered signals from the atmosphere are gathered by a 0.6 m telescope and separated into 10 optical channels using both dichroic beam splitters (for the main Raman and aerosol parameters) and a fiber optic (for the rotational Raman temperature channels). Both analog-to-digital and photon counting data acquisition are used and the minimum range resolution of the system is 7.5 m. To maximize the signal-to-noise of the water vapor measurement in the dry UTLS region, in addition to the cooled water vapor PMT, narrow telescope field-of-view (250 μ radians FWHM) and water vapor interference filter (0.25 nm) are used.

During MOHAVE-2009, profiles were acquired every 2 min to capture the evolution of the highly variable water vapor environment at Table Mountain. Optical channels in use during MOHAVE-2009 included Raman water vapor, nitrogen, liquid/ice water (Wang et al., 2003), 3 different unpolarized aerosol channels, parallel and perpendicular polarization aerosol channels, and two rotational Raman channels similar to that described by Di Girolamo et al. (2004), except that in this new configuration a filter polychromator approach (Behrendt and Reichardt, 2000) is used and the Stokes part of the spectrum is sampled as opposed to the anti-Stokes as in Di Girolamo et al. (2004). The quick look ALVICE temperature data have been compared with the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) measurements made during MOHAVE-2009 (Stiller et al., 2012). The comparison indicated that MIPAS temperatures are 1–2 K higher than lidar temperatures below the tropopause and lower than lidar temperatures above the tropopause. This same relationship was found in other comparisons (Stiller et al., 2012).

The Raman lidar contained in ALVICE operated on 13 nights during MOHAVE-2009, acquiring approximately 88 h of measurements of water vapor, aerosols, clouds and temperature over the period of 12–27 October 2009.

2.2 ALVICE ancillary instrumentation and measurements

MOHAVE-2009 presented the first opportunity to deploy all of the ALVICE instrumentation during a field campaign. All balloon preparations and launches for both Vaisala RS92 and cryogenic frost point hygrometer (CFH) (Vömel et al., 2007a) were led by Table Mountain Facility (TMF) staff, but several balloons included dual RS92 launches where the second RS92 was prepared by Goddard Space Flight Center (GSFC) staff and the data were received and recorded using the ALVICE ground stations. These simultaneous RS92 flights permitted the study of radiosonde data consistency and production variability. There were also several launches of the NOAA frost point hygrometer (Hurst et al., 2011b). The frost point hygrometer measures the frost point temperature of the air based on the chilled-mirror principle (Mastenbrook and Dinger, 1961), and uses a cryogenic heat sink for fast response and accurate measurement under very dry conditions (Vömel et al., 2007a; Hurst et al., 2011b). The frost point measurements are converted to RH and mixing ratio using temperature (T) and pressure (P) measured by the RS92, respectively, with an estimated total uncertainty of $\pm 4\%$ at the surface increasing to $\pm 10\%$ in the lower stratosphere. An extensive discussion of the radiosonde and frost point measurements during MOHAVE-2009 is given by Hurst et al. (2011b). Henceforth, the CFH and NOAA frost points will be considered equivalently, as done by Hurst et al. (2011b), and referred to as FP.

THPref

The ALVICE mobile trailer includes a surface reference system referred to as THPref (Temperature-Humidity-Pressure Reference) that provides NIST traceable measurements of temperature, relative humidity and pressure and contains a ventilated chamber for characterizing radiosonde accuracy prior to launch. The time series of surface data and derived products is shown in Fig. 1 where the SuomiNet total column water data are reported as well. Appendix A describes corrections applied to the RS92 RH data and the estimated total uncertainty of corrected RS92 measurements during MOHAVE-2009. For conditions of $\text{RH} > 10\%$, the total uncertainty is stated as $\pm (5\% + 0.5\% \text{RH})$, meaning 5 % of the measured RH value plus a 0.5 % RH component that is an increasingly significant percentage uncertainty as the RH decreases (e.g., at 50 % RH the uncertainty is $0.05 \times 50 + 0.5 = 3.0\% \text{RH}$ or $\pm 6\%$ of the 50 % RH value, and at 10 % RH the uncertainty is $\pm 10\%$). For conditions of $\text{RH} < 10\%$ the total uncertainty is $\pm (7\% + 0.5\% \text{RH})$ (e.g., at 3 % RH the uncertainty is $\pm 24\%$). These error assessments for the corrected data are altitude independent. By contrast, the error characteristics of uncorrected RS92 data have a significant altitude dependence (Miloshevich et al., 2009).

3 Previous ALVICE lidar measurements and contamination during MOHAVE-2009

Prior to MOHAVE-2009, the ALVICE system was deployed to the Howard University Beltsville Campus in Beltsville, MD, for the WAVES_2009 campaign in January–April 2009. During that campaign, long duration nighttime measurements of water vapor mixing ratios were made using the ALVICE lidar system. Four frost point hygrometers were launched during this campaign. Good agreement was found between ALVICE lidar, frost point hygrometer and available MLS data in the lower stratosphere. There was no indication of a moist bias in the ALVICE lidar UTLS water vapor measurements during WAVES_2009 and therefore we believe that the ALVICE lidar arrived for MOHAVE-2009 free of any significant fluorescence. The measurements during WAVES_2009 were made with the telescope protected from the outside environment through the use of an optical window. This window was removed at the beginning of MOHAVE-2009 to reproduce the measurement configuration of the ALVICE system during MOHAVE-II in 2007 (Whiteman et al., 2010) where the window was not used so as to maximize the signal returns.

The weather at Table Mountain during MOHAVE-2009, however, presented some quite different conditions than encountered during MOHAVE-II. There were periods of warm weather during which significant populations of flying insects were present in the vicinity of the lidar systems. On the night of 16 October, a large number of insects was attracted

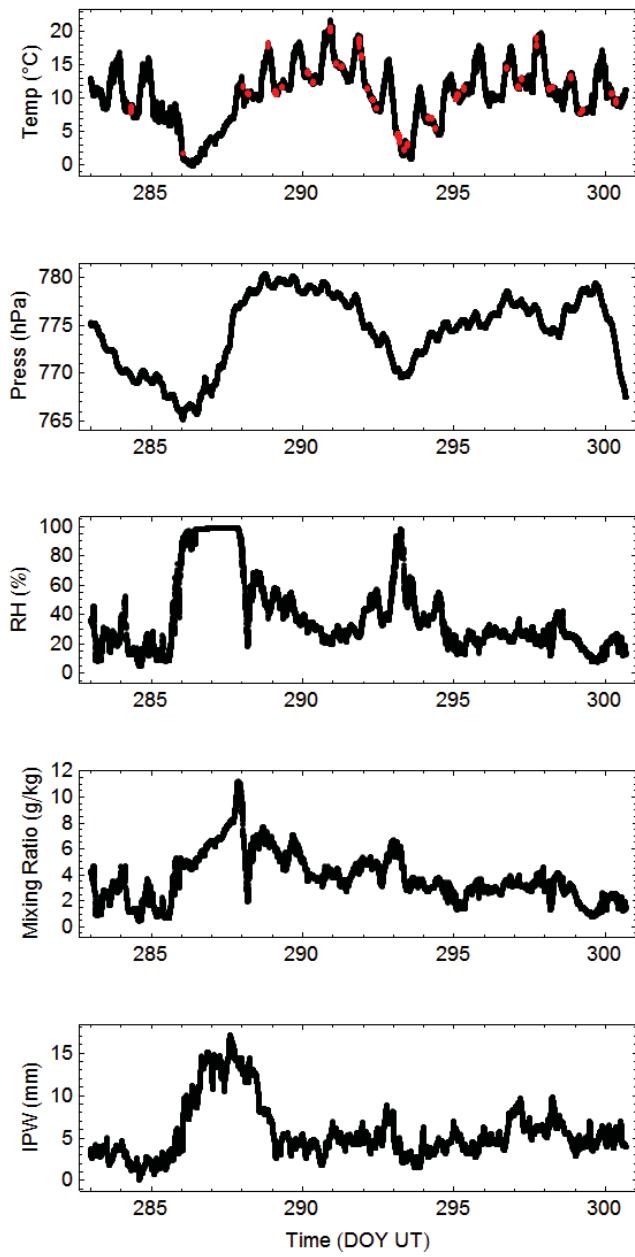


Fig. 1. Time series during MOHAVE-2009 of surface T , P , RH, mixing ratio and total column water from the THPRef and SuomiNet instruments that are part of the ALVICE instrumentation. The times when radiosondes were inserted into the ventilated chamber for comparative measurements are marked with red dots.

to the UV laser beam, which is emitted along the optical axis of the telescope. Insects that were attracted to the UV laser radiation were able to follow the laser beam into the trailer and enter both the telescope and the laser. Signals were lost during this data acquisition session due to insects obscuring the field stop of the telescope; thus, operations were curtailed early on this night. Considerable effort was expended in cleaning the laser and the telescope and in re-installing

the protective window that had been used in WAVES_2009 over the telescope. A proper cleaning of the ALVICE telescope receiver, however, requires an effort that was deemed too complicated and time consuming to attempt during a field campaign. Therefore, a noticeable amount of insect residue remained on the telescope primary and other components of the receiving optics after this cleaning. This was the configuration of the lidar receiver optical system for the measurements analyzed here. As will be described below, excess signal was found in the water vapor channel and influenced the UTLS portion of the profile. Some insects are known to have a strong fluorescence signature when stimulated with UV radiation (Byrdegaard et al., 2009). Therefore, it is believed that this additional signal was due to the fluorescence of the insect residue that remained deposited on the lidar telescope.

In the analysis that follows, the additional signal in the water vapor channel due to fluorescence is considered to be constant during the MOHAVE campaign for the following reasons. After the insect incident on 16 October, a window was installed over the ALVICE telescope. This window prevented further contamination inside the ALVICE trailer. Also, a large industrial fan (0.9 m diameter, $340\,000\,\text{l min}^{-1}$) was mounted on top of the trailer and blew across the window during lidar operation periods to prevent any significant material buildup on the window. In addition, the window was washed prior to every subsequent operation period during MOHAVE. Finally, as we will point out later in Sect. 6, the analysis of the data does not indicate significant variation of the correction value during the MOHAVE campaign.

4 The challenge of UTLS water vapor measurements using Raman lidar

The literature and other analyses reveal that there are three general causes of wet biases in Raman lidar UT water vapor measurements. The apparent excess amount of water vapor can be due to (1) the lidar instrument itself, (2) the data processing or (3) atmospheric contaminants. All three of these have played a role in the evolution of instrumental development and data processing that occurred during the 3 MOHAVE experiments in 2006, 2007 and 2009 where investigation of wet biases due to fluorescence of optical materials was a concern. Here we review these three sources of wet bias and in the following section we develop a mathematical model for correcting some of them. Following development of the model, lower stratospheric water vapor climatology is investigated as a means to help detect biases and correct for them.

4.1 Biases due to instrumental effects

There are earlier studies of upper tropospheric and lower stratospheric water vapor measurements using Raman lidar that have shown a wet bias. Sherlock et al. (1999) noted

excess water vapor signal corresponding to a wet bias of approximately 80 ppmv due to fluorescence in their fiber optic cable. Replacement of this fiber by an OH-rich fiber reduced the fluorescence contamination to 8 ppmv or less. However, the authors state that "... although here absorption of the elastic-backscatter signal occurs in the fiber-optic cable used for signal transfer, it could arise in any optical component. Thus, fluorescence processes are a potential source of systematic error in Raman Stokes measurements, particularly in the case of water vapor, where the signals are weak and susceptible to contamination."

Insufficient blocking of unwanted, out-of-band signals in the water vapor channel can also create a spurious wet bias in the water vapor mixing ratio measurement that has similar appearance to the fluorescence contamination just mentioned. Likewise, the presence of afterpulsing (Piironen, 1994) in the water vapor PMT could create an apparent wet bias. In the case of afterpulsing the magnitude of the wet bias at a given altitude is related to the amount of water vapor signal that the PMT has produced at lower altitudes.

4.2 Biases due to data processing

Ferrare et al. (2004) discuss the Atmospheric Radiation Measurement (ARM)–First International Satellite Cloud Climatology Project (ISCCP) Regional Experiment (FIRE) Water Vapor Experiment (AFWEX) field campaign that took place in December 2000 at the Department of Energy's Southern Great Plains (SGP) ARM site. The initial results showed the Cloud and Radiation Testbed (CART) Raman lidar to possess a wet bias in the upper troposphere with respect to the reference airborne water vapor instrumentation. The CART Raman lidar is calibrated based on a total column water comparison with the ARM microwave radiometer, so the wet bias was found to be due to an error in the overlap correction used in the lidar data processing (Ferrare et al., 2004) that influenced the magnitude of the total column water calculated from lidar.

Whiteman et al. (2006c) report on the results of the AWEX-G (AIRS Water Vapor Experiment-Ground) field campaign held at the same SGP site in 2003. This campaign was staged due to the early results of the NASA Aqua satellite validation program that indicated significant disagreement between Raman lidar and Vaisala radiosonde measurements of water vapor. Water vapor profiles from both Raman lidar and Vaisala RS90 were being used as sources of calibration/validation data in AIRS fast forward model studies. The simulated radiances calculated using the lidar and radiosonde profiles in the AIRS forward model implied a disagreement of approximately 25 % in the upper tropospheric water vapor measurements between the scanning Raman lidar and the Vaisala radiosondes. A Raman lidar UT wet bias of approximately 12 % was found. A dry bias in the radiosondes was also quantified (Miloshevich et al., 2006). The lidar wet bias was corrected by accounting for the temperature-dependence

of Raman scattering and improving the lidar overlap function (Whiteman et al., 2006c). The Vaisala dry bias was addressed by empirical correction (Miloshevich et al., 2006) of the RS90 to obtain best agreement with frost point hygrometer. After applying these independently determined corrections, the scanning Raman lidar measurements and those of empirically corrected Vaisala radiosonde were found to agree at the 5 % level in the upper troposphere.

4.3 Biases due to atmospheric constituents

In addition to the fluorescence of atmospheric insect remains that led to the moist bias in the ALVICE Raman lidar data reported here, there is other evidence in the literature of direct fluorescence of atmospheric particles presenting a contaminating signal to Raman water vapor lidars. Immler et al. (2005) detected enhancements in their water vapor signal due to the fluorescence of lower stratospheric aerosols. Furthermore, aerosols of biological origin such as bacteria, spores or pollen and leaf matter are found in the atmosphere and can present a strong fluorescence signature that, due to the spectral nature of the fluorescence, can create an apparent wet bias in the water vapor measurement (Saito et al., 2010). The fluorescence signatures of individual aerosol particles in the ambient atmosphere have been studied, showing that a small percentage of these particles possesses a strong fluorescence signature (Pan et al., 1999; Pinnick et al., 2004) in the spectral region where Raman water vapor lidar measurements are made. These results confirm earlier work (Gelbwachs and Birnbaum, 1973) indicating that aerosol fluorescence posed potential problems for Raman lidar measurements of trace gases.

4.4 Wet biases in other lidar systems

Study of the data from well-known Raman lidar systems that did not participate in the MOHAVE experiments indicates the presence of wet UT biases in these systems as well. For example, 5 yr of CART Raman lidar (Turner et al., 2002; Ferrare et al., 2004) data have been studied covering the period of 1998–2003. A persistent wet bias of approximately 8–12 ppmv was found in the lower stratosphere. Approximately 10 yr of data from the Tor-Vergata lidar outside of Rome (Dionisi et al., 2009) have been studied, with the conclusion that a persistent wet bias of approximately 10 ppmv in the lower stratosphere exists as well (Gian Luigi Liberti, personal communication, June 2011). Selected data from ten years of measurements with the Meteorological Research Institute Raman lidar in Tsukuba, Japan (Sakai et al., 2007) have also been studied, revealing a persistent wet LS bias of approximately 50 ppmv (Tetsu Sakai, personal communication, May 2011). These biases are large with respect to typical lower stratospheric water vapor amounts of a few ppmv but small with respect to tropospheric values that can exceed 10 000 ppmv.

Furthermore, it should be noted that the wet bias present in the CART Raman lidar data record between 1998–2003 encompasses the period of the AFWEX experiment (Ferrare et al., 2004), during which the finally processed CART Raman lidar data agreed well with the reference in the UT. It may be that the focus of the UT investigations during the AFWEX mission, which was on the integrated water between 7 and 12 km, masked the presence of this small moist bias that manifested itself primarily at altitudes above 12 km. Fortunately, the complete raw data from the lidar are archived within the DOE/ARM program permitting this question to be studied further.

The details above indicate that wet UTLS biases in Raman lidar data are rather common and derive from various sources, including atmospheric, instrumental and algorithmic effects. The apparent frequent occurrence of these wet biases can be related to the very large challenge of measuring the low values of water vapor mixing ratio in the UTLS using the Raman lidar technique. Recent ALVICE Raman lidar measurements acquired in the Greenbelt, MD, vicinity indicate that the signal-to-noise ratio of the water vapor measurements at an altitude of 13 km m.s.l. is approximately the same as the signal-to-noise ratio of the measurements of Raman nitrogen at 60 km or the Rayleigh signal at 80 km. In other words, because of the small water vapor amounts in the UTLS, very low signal-to-noise ratios exist and the potential for small error sources to bias the data is thus magnified.

The main point to be made concerning the preceding discussion is that the potential of a bias being present or developing in a Raman lidar measurement of UTLS water vapor seems too high to not be concerned with on a routine basis. Correcting systematic effects such as these is consistent with the recommendations of the Joint Committee on Guides in Metrology in their Guide to the Expression of Uncertainty in Measurements (GUM) (JCGM/GUM, 2008), where it is stated that “Systematic error, like random error, cannot be eliminated but it too can often be reduced. If a systematic error arises from a recognized effect of an influence quantity on a measurement result, hereafter termed a systematic effect, the effect can be quantified and, if it is significant in size relative to the required accuracy of the measurement, a correction ... or correction factor ... can be applied to compensate for the effect. It is assumed that, after correction, the expectation or expected value of the error arising from a systematic effect is zero.” Also stated is that “It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects.” Following these recommendations, we first present a technique for correcting certain wet biases and then consider methods for assessing the presence of wet biases in the UTLS and deriving the necessary correction.

5 Mathematical model for signal-dependent water vapor bias

The following model is developed to address the situation where additional signal is present in the water vapor channel of a Raman lidar and that additional signal is proportional to the magnitude of the signal backscattered at the laser wavelength. This situation could be created by fluorescence of optical materials in the lidar receiver or insufficient blocking of any of the Rayleigh-Mie, Raman N₂ or O₂ signals in the Raman water vapor channel. The case of excess signal in the water vapor channel due to insufficient blocking of the Rayleigh-Mie signal or fluorescence will be handled here. The case for leakage of the Raman nitrogen or oxygen return is discussed in Sect. 5.1. Note that a situation of excess signal due to fluorescence of airborne particles as described in Immler et al. (2005) or of afterpulsing of the PMTs (Piironen, 1994) would not be addressed by this correction technique. Correction for fluorescence of airborne particles would require a knowledge of the intensity of fluorescence and its spectrum. This information is generally not available. Corrections for afterpulsing effects have been performed (Piironen, 1994) but the mathematical formulation differs from that developed here for signal-dependent biases, since it is the intensity of the water vapor signal (as opposed to the Rayleigh-Mie one) that stimulates additional signal in the water vapor channel, and the magnitude of the excess signal at a particular height in the profile is related to the total water vapor content below that altitude.

We will use the formulation of the lidar equations used by Whiteman (2003a,b). The water vapor mass mixing ratio may be expressed as

$$w = k \frac{O_N(r)}{O_H(r)} \frac{F_N(T(r))}{F_H(T(r))} \frac{P(\lambda_H, r)}{P(\lambda_N, r)} \frac{\frac{d\sigma_N(\pi)}{d\Omega}}{\frac{d\sigma_H(\pi)}{d\Omega}} \frac{\xi(\lambda_N)}{\xi(\lambda_H)} \Delta \tau(\lambda_N, \lambda_H, r), \quad (1)$$

where the constant of proportionality, k , has a value of approximately 0.486 and is determined by the molecular weights of water vapor and nitrogen and the fractional abundance of nitrogen in the atmosphere (Whiteman, 2003b). $P(\lambda_X, r)$ is the backscattered power (after subtracting any background contribution due, for example, to skylight or detector noise) received at λ_X due to Raman scattering from either water vapor (H) or nitrogen (N) as a function of range, r . $O_X(r)$ is the overlap function for either Raman channel. $\xi(\lambda_X)$ is the total lidar receiver optical efficiency at the wavelengths of the water vapor or nitrogen signals and includes factors such as the reflectivity of the telescope, the transmission of conditioning optics, the transmission of any filters and the quantum efficiency of the detector. $\frac{d\sigma_X(\pi)}{d\Omega}$ is shorthand notation for the Raman backscatter cross section when stimulated at the laser wavelength. $F_X(T(r))$ is a factor that carries all the temperature-dependence of the Raman scattering

process and thus appears as a multiplier of the traditional Raman lidar equation (Whiteman, 2003a). $\Delta\tau(\lambda_N, \lambda_H, r)$ is the “differential transmission” between the λ_N and λ_H and is calculated typically from a radiosonde measurement of atmospheric density as follows:

$$\Delta\tau(\lambda_N, \lambda_H, r) = \exp\left(-\int_0^r \{\alpha(\lambda_N, r') - \alpha(\lambda_H, r')\} dr'\right) \quad (2)$$

where $\alpha(\lambda_X, r)$ is the total extinction coefficient at the pertinent wavelength due to scattering and absorption by molecules, particles and any other atmospheric constituents such as water droplets or ice crystals as a function of range along the path of the laser beam.

Consider now the situation where there is excess signal in the water vapor channel due to a signal-dependent source such as interference filter bleed-through or fluorescence. This dependence is expressed as $\xi_0 P(\lambda_R, r)$ in the following equations, where ξ_0 is the ratio of the additional, spurious signal in the water vapor channel and the Rayleigh-Mie signal at the laser wavelength, which is represented by $P(\lambda_R, r)$. We will assume that the magnitude of any such signal-dependent bias in the nitrogen channel is negligible in the ratio shown in Eq. (1). This assumption is based on the typical situation where the Raman nitrogen signal strength is at least 6 orders of magnitude larger than that from water vapor in the upper troposphere, below which the signal-dependent biases considered here become increasingly less significant due to the generally much larger water vapor mixing ratios at lower altitudes.

If the water vapor channel is contaminated by an amount proportional to the backscattered Rayleigh-Mie signal intensity, the apparent water vapor mixing ratio, w^* , can be expressed as

$$w^* = k \frac{O_N(r)}{O_H(r)} \frac{F_N(T(r))}{F_H(T(r))} \frac{P(\lambda_H, r) + \xi_0 P(\lambda_R, r)}{P(\lambda_N, r)} \quad (3)$$

$$= k \frac{\frac{d\sigma_N(\pi)}{d\Omega}}{\frac{d\sigma_H(\pi)}{d\Omega}} \frac{\xi(\lambda_N)}{\xi(\lambda_H)} \times \Delta\tau(\lambda_N, \lambda_H, r) \quad (4)$$

$$= k \frac{O_N(r)}{O_H(r)} \frac{F_N(T(r))}{F_H(T(r))} \frac{P(\lambda_H, r)}{P(\lambda_N, r)} \frac{\frac{d\sigma_N(\pi)}{d\Omega}}{\frac{d\sigma_H(\pi)}{d\Omega}} \frac{\xi(\lambda_N)}{\xi(\lambda_H)} \times \Delta\tau(\lambda_N, \lambda_H, r) \left(1 + \frac{\xi_0 P(\lambda_R, r)}{P(\lambda_H, r)}\right) \quad (5)$$

$$= w \left(1 + \frac{\xi_0 P(\lambda_R, r)}{P(\lambda_H, r)}\right) \quad (5)$$

$$w = w^* \frac{P(\lambda_H, r)}{P(\lambda_H, r) + \xi_0 P(\lambda_R, r)}, \quad (6)$$

where the value of ξ_0 can be determined by comparison of the lidar mixing ratio profile with a reference profile. Equation (6) gives an exact solution when the value of ξ_0 is known. Under certain assumptions the correction for the excess signal can be approximated by a simple subtraction in mixing ratio space, as will now be shown.

Working with Eq. (5) and assuming that no significant aerosol scattering is present where the contamination is significant, generally in the upper troposphere and lower stratosphere, the Raman nitrogen signal is proportional to the Rayleigh signal after correction for the differential transmission of the two signals and accounting for the temperature dependence of Raman scattering. This assumption of insignificant aerosol scattering was met during MOHAVE-2009. In the presence of significant aerosol scattering, the full form of the correction given in Eq. (6) would need to be used. Assuming that no significant aerosol scattering is present and noting that $P(\lambda_R, r)/P(\lambda_N, r) \propto F_R(T(r))/F_N(T(r))$ $\Delta\tau(\lambda_R, \lambda_N, r)$ and that $P(\lambda_N, r)/P(\lambda_H, r) \propto F_N(T(r))/F_H(T(r))$ $\Delta\tau(\lambda_N, \lambda_H, r)$ results in the following expression:

$$w^* = w \left(1 + \xi_1 \frac{F_R(T(r))}{F_H(T(r))} \frac{\Delta\tau(\lambda_R, \lambda_H, r)}{w}\right), \quad (7)$$

where a new constant ξ_1 has been introduced to account for the ratio of the signal backscattered at the laser wavelength and the Raman nitrogen signal. Note also that this differential transmission term $\Delta\tau(\lambda_R, \lambda_H, r)$ differs from the earlier defined one of $\Delta\tau(\lambda_N, \lambda_H, r)$ used in Eq. (1). Equation (7) has solution of

$$w = w^* - \xi_1 \frac{F_R(T(r))}{F_H(T(r))} \Delta\tau(\lambda_R, \lambda_H, r) \quad (8)$$

$$\sigma_w^2 \approx \sigma_{w^*}^2 + \sigma_{\xi_1}^2. \quad (9)$$

From Eq. (8) it can be seen that, with the assumption described above, the corrected mixing ratio is calculated from the measured mixing ratio by subtracting a constant times the differential transmission profile $\Delta\tau(\lambda_R, \lambda_H, r)$ multiplied by the temperature-dependence ratio. This ratio, $F_R(T(r))/F_H(T(r))$, has little practical effect since the correction is generally significant only in the upper troposphere and lower stratosphere where the temperature profile first decreases to the tropopause and then starts to increase, thus minimizing the overall change in temperature. For simplicity, therefore, the temperature-dependence of the correction can be omitted. The uncertainty in the corrected mixing ratio is also given in Eq. (9), where σ_x^2 refers to the variance of the respective quantities. For this approximation, the contribution to the uncertainty budget due to both the differential transmission and temperature-dependence factors, which typically have relative uncertainties of less than 1 % each, have been excluded. Also, although the term ξ_1 is a constant in the model, there is uncertainty attributed to the determination of the value of the constant. In the analysis presented here, the uncertainty in ξ_1 is taken to be the standard error of the mean differences between the corrected ALVICE profiles and the FP profiles and, as will be shown later in Sect. 6, is estimated to be less than 0.25 ppmv during MOHAVE-2009 campaign.

The correction described here is generally only significant in the upper troposphere and above, a range over which

the differential transmission between the laser wavelength and the water vapor wavelength due to molecular scattering changes by approximately 5 % and where the temperature-dependent ratio changes by less than 1 %. Therefore, the uncertainty introduced by subtracting a constant from the contaminated mixing ratio as an alternative to Eq. (8) is approximately $\pm 2.5\%$. That being the case, alternate approximations for the corrected water vapor mixing ratio and its variance are given by

$$w = w^* - \xi_2 \quad (10)$$

$$\sigma_w^2 \approx \sigma_{w^*}^2 + \sigma_{\xi_2}^2 \quad (11)$$

where the range-dependent differential transmission term has been dropped, a new correction term, $\xi_2 \approx \xi_1 F_R(T(r))/F_H(T(r)) \Delta \tau(\lambda_R, \lambda_H, r)$, has been introduced and the uncertainty in w is similar to that given above. The three forms of the corrected mixing ratio equations derived here are compared in Sect. 5.2 after addressing the case of leakage of one of the Raman signals into the water vapor channel.

It should be noted that fluorescence is modeled here as being instantaneous. Laboratory studies of reference dyes indicate lifetimes generally of less than 10 ns (Boens et al., 2007), while studies of polycyclic aromatic hydrocarbons indicate lifetimes generally of 20–30 ns, with maximum values of 36 ns (Dvorak et al., 1997). When sensed by lidar, the non-zero lifetime of the fluorescence would delay and stretch the return signals. However, the correction developed here is only significant in the upper troposphere and lower stratosphere where the vertical resolution of the lidar data ranges between 500 and 750 m. This vertical resolution corresponds to temporal bin widths between 5000 to 7500 ns (5–7.5 μ s). The delay in the fluorescence signal due to this lifetime is considered to be insignificant when compared with the vertical resolution of the lidar data. This is the reason that the fluorescence is modeled as being instantaneous.

5.1 Excess signal due to Raman signal leakage

A similar situation that should be considered in this context is created by an insufficient optical density in the water vapor filter at the wavelength of the Raman nitrogen (or oxygen) return. This becomes particularly important in the measurement of upper tropospheric and lower stratospheric water vapor because of the very low amount of water vapor and the relatively higher amount of molecular nitrogen (or oxygen). Considering that a typical LS mixing ratio might be 5 ppmv and that of molecular nitrogen $\sim 780\,000$ ppmv, in order to limit the contamination due to Raman nitrogen to less than 1 % of the water vapor signal, the blocking required in the water vapor filter at the Raman nitrogen wavelength is approximately 10^7 . If the filter does not have sufficient blocking of the Raman nitrogen signal, this will appear as a wet bias in the water vapor channel, similar to fluorescence or bleed-through at the laser wavelength. In this case, using a

similar derivation as above, the correction equation and its variance are found to be

$$w = w^* - \xi_3 \Delta \tau(\lambda_N, \lambda_H, r) \quad (12)$$

$$\sigma_w^2 \approx \sigma_{w^*}^2 + \sigma_{\xi_3}^2 \quad (13)$$

where ξ_3 represents the leakage of the water vapor filter at the Raman nitrogen wavelength and for simplicity the temperature-dependence factors have been omitted. Considering now the relative magnitude of $\Delta \tau(\lambda_N, \lambda_H, r)$ and $\Delta \tau(\lambda_R, \lambda_H, r)$, the following equations pertain with an uncertainty of less than $\pm 2.0\%$:

$$w = w^* - \xi_4 \quad (14)$$

$$\sigma_w^2 \approx \sigma_{w^*}^2 + \sigma_{\xi_4}^2 \quad (15)$$

where $\xi_4 \approx \xi_3 \Delta \tau(\lambda_N, \lambda_H, r)$. The case for leakage at the Raman oxygen signal is handled in a similar manner. The conclusion is that, to a very good approximation under many upper tropospheric and lower stratospheric conditions, excess water vapor signal due to bleed-through of the filter (at the laser fundamental or the Raman shifted wavelengths for nitrogen or oxygen) or due to fluorescence in the telescope receiver can be corrected quite accurately by the subtraction of a constant from the mixing ratio profile.

5.2 Comparison of the correction equations

Equations (6), (8) and (10) were evaluated using several long-duration data runs acquired during the MOHAVE-2009 campaign. ALVICE Raman lidar measurements were used on the nights of 20, 21, 25, and 27 October to compare the effect of the different correction equations. On these nights there were a total of 5 FP measurements. The magnitude of the correction constants ξ_0 , ξ_1 and ξ_2 were chosen empirically to yield best mean agreement with the FP in the 10–20 km altitude range. Other more sophisticated methods of determining the correction value were considered but yielded no significant difference from this simple approach. The mean comparisons of the uncorrected lidar data and the corrected data using the three correction techniques are shown in the left panel of Fig. 2. The results using Eqs. (6), (8) and (10) are referred to as “Corr1”, “Corr2” and “Corr3”, respectively. The two approximate corrections shown in green and blue agree within 1 % of each other and are always in agreement with the exact correction to better than 5 % (generally much better), as expected.

It would be preferred to use the exact form of the correction given by Eq. (6). This was attempted using a single hour of lidar data from the night of 26 October and the results are shown in the right panel of Fig. 2. Here, the signal-to-noise in the Rayleigh-Mie signal averaged over 1 h is insufficient to produce a stable correction. The instability is due to the presence of the correction term $\xi_0 P(\lambda_R, r)$ in the denominator of the equation, the generally small values of $P(\lambda_H, r)$ and the lower signal-to-noise in the Rayleigh-Mie data for the

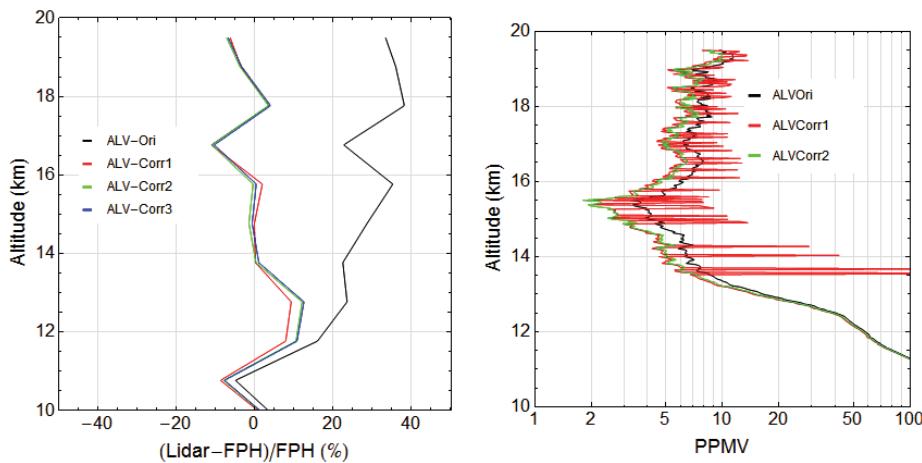


Fig. 2. Left panel: mean comparison of three correction equations using long duration measurements on 4 nights. Right panel: comparison of the the first two correction techniques using a single hour of lidar data.

one-hour measurements versus the all-night ones shown on the left. For this reason, Eq. (8) was chosen as the preferred correction technique to use for processing the MOHAVE-2009 data.

Having determined that Eq. (8) was the preferred correction technique to use, a single value of ζ_1 was determined for the entire MOHAVE campaign using all available lidar/FP comparisons over the course of the MOHAVE campaign. All ALVICE water vapor data were corrected using a fixed value of $\zeta_1 = 1.5 \text{ ppmv}$.

6 Water vapor profile and total column water vapor comparisons

Comparisons of the fully processed water vapor profiles from the ALVICE Raman lidar were made with respect to RS92 and FP. These comparisons were used both for calibration of the Raman lidar water vapor measurements using a new, adaptive technique that is described in Appendix A1 as well as to validate the profiling performance of the lidar. As part of this validation, the mean normalized differences of the 1 h lidar summed profiles were formed with respect to Vaisala RS92 and FP using 1 km layer averages. The mean normalized difference of the “all-night” summed ALVICE lidar profiles were also formed with respect to FP. These three comparisons are shown in Fig. 3. There were 19 comparisons of ALVICE lidar and RS92 and 11 comparisons of ALVICE lidar and FP. The plot of 1 h lidar measurements versus RS92 shows a moist bias of the lidar of approximately 20% in the lowest 0.5 km. The study of radiosonde profile similarity performed in Appendix A2 indicates that some of this bias is due to a persistent tendency for the atmospheric layer directly above the mountaintop lidar site to be moister than the layer immediately downwind of the site, where the radiosonde sampled. Above this surface layer, the ALVICE

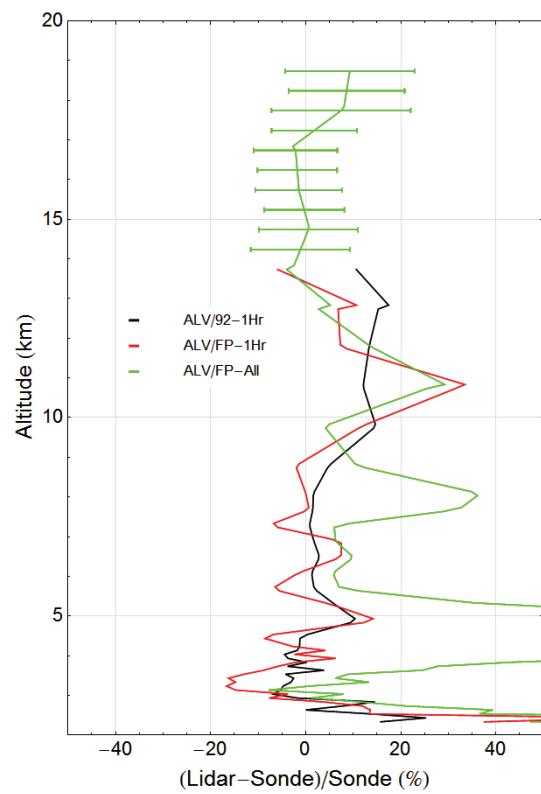


Fig. 3. Mean normalized differences of all available 1 h and all-night sum ALVICE profiles with RS92 and FP. The standard deviation of the differences is plotted above 14 km for the FP all-night lidar comparison. See text for details.

lidar 1 h profiles agree with the RS92 within 10–15% up to an altitude of 14 km, with the radiosonde profiles showing a tendency to become increasingly dry with respect to the lidar profiles for altitudes above approximately 8 km. The analysis of RS92 and FP data shown in Appendix A also

indicates that the corrected RS92 becomes increasingly dry with respect to the FP with increasing altitude. This behavior is consistent with the relative behavior of the lidar and RS92 profiles. The differences between RS92 and FP or lidar are within the uncertainty of the Vaisala correction algorithm but further support the possibility that the RS92 calibration may have changed since the time the RS92 correction algorithm was determined. These points are further discussed in Appendix A.

The comparison of 1 h lidar profiles and the frost point measurements shown in Fig. 3 generally displays more scatter than the comparison versus the RS92 due to the reduced number of comparisons. The feature in the range of 10–11 km that shows a large positive bias of approximately 30 % occurs where, in a small number of cases, the atmospheric profile of water vapor showed a rapid decrease above 10 km. It is believed that atmospheric variability between directly overhead (where the lidar was measuring) and where the balloon had drifted is the cause of this discrepancy.

The comparison of all-night lidar profiles and FP used the same 11 FP launches as in the 1 h comparisons. The tendency for the all-night lidar profiles to exhibit a moist bias mentioned in Appendix A1 with respect to Fig. A7 is evident here as well. Below approximately 12 km, the lidar profiles are wetter than the FP profiles by approximately 10 %, discounting the regions of high atmospheric variability at altitudes of approximately 4, 8 and 11 km where deviations were higher. The differences at 4, 8 and 11 km are believed to be mainly due to atmospheric variability and not due to real instrument measurement differences. Variability in the troposphere is frequently an issue when comparing a small number of profiles not acquired from the same atmospheric volume. Above 12 km, where the measurements of FP may be expected to generally be more representative of the all-night lidar measurements, the comparisons improve considerably, with ALVICE lidar agreeing with FP within generally better than 10 % to beyond 18 km.

The question of whether there was significant variation in the value of the correction factor, ζ_1 , during the MOHAVE mission can be addressed with the aid of Fig. 3 where the standard deviation of the comparisons of all FP and all-night summation lidar comparisons is plotted above 14 km. Consider that the displayed standard deviations of the normalized differences between corrected ALVICE and FP range between approximately 10 and 15 % (standard error of 3–5 % or < 0.25 ppmv). The standard deviation of the lidar data alone ranged from 8 to 15 % in this altitude range, while the accuracy of the FP is given as 10 % in this part of its profile. There is also variability contributed by the spatiotemporal mismatch between lidar and FP measurements, as well as potentially by variation in the correction factor. The entire standard deviation budget is likely accounted for, however, by considering only the variability in the lidar and the FP without considering any contribution due to spatiotemporal mismatch or variation in the correction factor. It seems

reasonable to state, therefore, that there is no evidence for significant day-to-day variability in the correction factor, ζ_1 , during the MOHAVE campaign. This result indicates that the use of an optical window over the lidar telescope augmented by a large capacity fan and daily window washings may be sufficient measures for preventing the accumulation of fluorescing material on a Raman lidar system.

7 Quality control of upper tropospheric Raman water vapor lidar measurements

Global climate models indicate that upper tropospheric water vapor is anticipated to increase significantly during the current century (Soden et al., 2005; Held and Soden, 2006; Oman et al., 2008; Boers and Meijgaard, 2009; Whiteman et al., 2011b). Water vapor mixing ratio increases of up to 1 % per year or higher are predicted by these models, with the largest trends anticipated in the tropical upper troposphere at the 150–250 hPa levels (Soden et al., 2005; Boers and Meijgaard, 2009; Whiteman et al., 2011b). Furthermore, recent studies indicate that these same pressure levels may be the most efficient ones for monitoring long-term trends in water vapor (Boers and Meijgaard, 2009; Whiteman et al., 2011b). By contrast, lower stratospheric water vapor is anticipated to increase less (Oman et al., 2008; Whiteman et al., 2011b). The variability in stratospheric water vapor is also substantially lower than in the upper troposphere, as evidenced by the 30-yr Boulder time series from NOAA FPH (Hurst et al., 2011a). This implies that monitoring trends in the lower stratosphere requires much more precise measurements than in the upper troposphere, where random uncertainties of 50 % and more appear tolerable (cf. Fig. 3 from Whiteman et al., 2011b). It also implies that additional noise or imperfections in the data record due to increases in random error, data gaps or biases will have a more significant effect on the time to detect trends in the lower stratosphere than in the upper troposphere. The error budget of a Raman lidar system increases with altitude so that the random errors are significantly larger in the LS than in the UT, implying that it may be more practical to consider Raman lidar for monitoring trends in the UT than the LS. The potential for wet biases being a larger fractional contamination in the LS than in the UT strengthens the argument that the focus of Raman water vapor lidar trend monitoring efforts should be in the UT and not the LS. Still, the lower stratospheric water vapor measurements of those Raman lidar systems able to measure that high into the atmosphere can be useful both for quality control of the data as well as a source of correction that improves the upper tropospheric water vapor measurements, potentially making them suitable for water vapor trend monitoring studies.

Consider Fig. 4, which presents the monthly average water vapor mixing ratios over Table Mountain, CA, as measured by the Microwave Limb Sounder (MLS) using v3.3

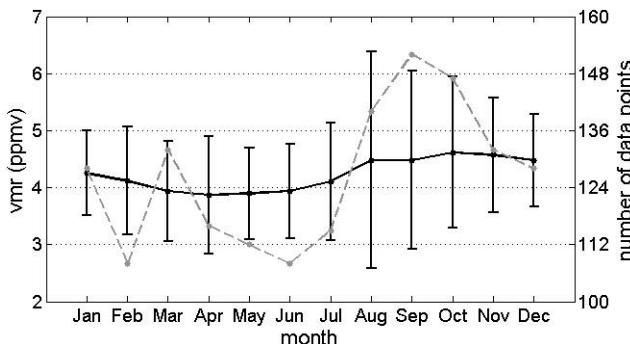


Fig. 4. Mean monthly water vapor mixing ratios (ppmv) measured by MLS over the altitude range of ~ 17 – 19.5 km. The average was computed using MLS v3.3 data from August 2004 until February 2011. The scale for the mixing ratio (solid black line) is on the left while the scale for the number of measurements used in each average (dashed grey line) is shown on the right. The error bars plotted are 2σ .

data from August 2004 to February 2011. The value plotted is the mean water vapor mixing ratio within the height range of 17 to 19.6 km. For the MLS data, geopotential altitudes are used. They differ by less than 1 % from the geometrical altitude in this height range. The number of retrievals averaged over the 7-yr period is plotted with a dashed line. The 2σ error bars plotted cover approximately 95 % of the number of cases. Thus, 95 % of the measurements at Table Mountain over this period of 7 yr fall within the range shown by the error bars on this plot. The absolute accuracy of MLS water vapor measurements at approximately 150 hPa is given (Livesey et al., 2011) as approximately 15 % with single-profile MLS precision reported as 0.2–0.3 ppmv (Lambert et al., 2007). Comparisons of MLS with CFH (Vömel et al., 2007b) above the tropical tropopause show agreement within 2–6 %. If we take, then, the results in Fig. 4 to represent a current climatology of lower stratospheric water vapor over Table Mountain accurate at the 5–15 % level, we can use it as a means to quality control Raman water vapor lidar data. For example, if Raman lidar LS measurements fall outside of the mean $\pm 2\sigma$ values shown in Fig. 4, then there is a 95 % likelihood that the Raman lidar data are biased in the lower stratosphere. This determination can be used as a source of correction to the data, following the guidelines of the GUM JCGM/GUM (2008) where systematic biases are assumed to be corrected. Recalling the many years of Raman lidar data described in Sect. 4.4 that indicate lower stratospheric water vapor mixing ratios outside of a reasonable range of values, this technique could also provide a way to correct those data. We will now test this concept on the ALVICE Raman lidar data and compare with the correction achieved by comparison with actual frost point measurements during the MOHAVE-2009 campaign.

During the MOHAVE-2009 campaign, 11 frost point hygrometer ascents (Hurst et al., 2011b) were available for

direct comparison with the ALVICE lidar profiles. This permitted the calculation of a correction value, ζ_1 , in Eq. (8) that gave good agreement with the mean frost point measurements in the altitude range of 10–20 km. As stated earlier, we believe the dominant cause of the wet bias present in the ALVICE lidar measurements to be fluorescence from insect material that came in contact with the telescope mirror early in the MOHAVE-2009 campaign. The value of $\zeta_1 \Delta \tau (\lambda_R, \lambda_H, r)$, therefore, can be considered to be a best estimate of the excess water vapor signal due to fluorescence from insect contamination for the configuration of the lidar during MOHAVE-2009. The difference between the mean lidar measurement and the MLS climatology can also be used to define ζ_1 or ζ_2 in Eqs. (8) and (10), respectively. Correction to a reasonable mean value in the LS would provide corrected data extending down into the upper troposphere where (1) trend detection is less influenced by additional noise sources and thus may be easier to perform, and (2) water vapor mixing ratios are much higher, thus decreasing (a) the relative measurement uncertainty of Raman lidar systems and (b) the magnitude of the influence of the assumed mean LS water vapor mixing ratio.

We now will use the MLS climatology shown in Fig. 4 to calculate the wet bias correction value, ζ_1 , and compare the results with the value calculated from direct comparison with the frost point reference measurements made during MOHAVE-2009. The mean mixing ratio measured by frost point during MOHAVE-2009 between the altitude ranges of 17.0 km and 19.6 km was approximately 4.7 ppmv, which is in very good agreement with the MLS mean climatology shown in Fig. 4. Determining the correction value, ζ_1 , for the ALVICE lidar from the MLS climatology between the altitude ranges of 17.0 km and 19.6 km, which is above the tropopause below which the MLS data have known biases (Livesey et al., 2011), yields a value of ζ_1 that is approximately 0.3 ppmv larger than the value achieved through comparison with frost point hygrometer between the ranges of 10 km to 20 km. ALVICE profiles using these two values of ζ_1 in Eq. (8) are included in Fig. 5.

In the left panel of the figure is shown the mean profiles for all available data during MOHAVE-2009 for FP, MLS and the ALVICE lidar, where both “best estimate” data, which is described in Appendix A4, and with correction derived from MLS climatology, as discussed above, are shown. The MLS vertical resolution in this altitude range is approximately 2.5 to 3.1 km (Livesey et al., 2011). For altitudes above 12 km, the ALVICE lidar profiles possess vertical resolution of approximately 0.5–0.75 km using the Verein Deutscher Ingeneure (VDI4210, 1999) definition of vertical resolution for lidar systems. All profiles generally reveal similar features, except for the departure of the MLS profile from the others below an altitude of approximately 14 km. Biases and deviations are better revealed by the plot in the right panel that shows the normalized differences, using 1 km layers, of MLS and ALVICE with respect to the

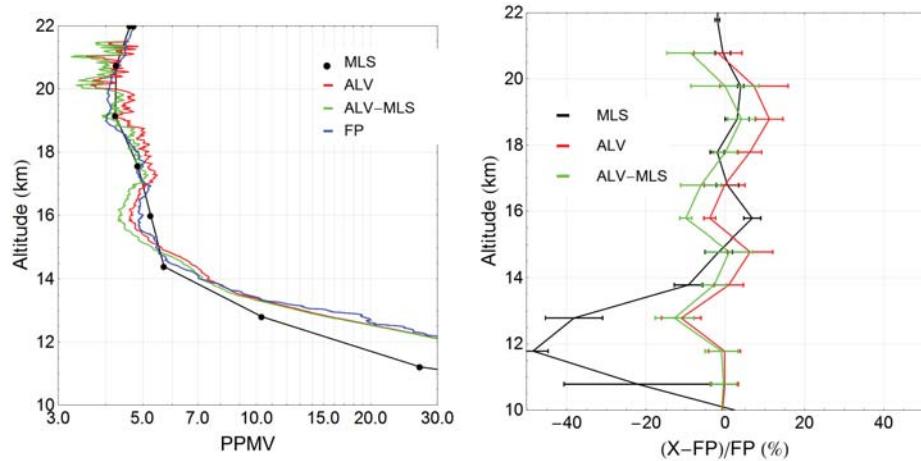


Fig. 5. Left panel: mean profiles of MLS, ALVICE (best estimate) and ALVICE (with correction based on MLS climatology) and FP for all data available during the MOHAVE campaign. Right panel: mean normalized differences in 1-km thick layers of MLS, ALV-best estimate, ALV-MLS corrected with frost point hygrometer using the same data as on the left.

FP profiles. MLS agrees with frost point hygrometer within generally better than 10 % above an altitude of 14 km. The previously mentioned upper tropospheric dry bias in MLS is evident below this altitude, reaching a maximum of approximately 40 % at an altitude of 12 km. The two versions of ALVICE profiles agree within 10 % of the frost point from 10 to 20 km.

Figure 5 indicates that similar results are obtained in the ALVICE data if MLS climatology is used to derive the signal-dependent bias correction for use in Eq. (8) instead of the actual frost point measurements made during MOHAVE-2009. This should not be surprising given the fact that the MOHAVE-2009 mission took place during more than half of the month of October and there were good statistics for UTLS measurements from both FP and ALVICE lidar during that period. Thus, the water vapor measurements made during MOHAVE-2009 are likely to represent a mean value for the month of October that is consistent with climatology.

As mentioned, recent results (Whiteman et al., 2011b) indicate that the most efficient level in the atmosphere for revealing trends in the atmospheric water vapor mixing ratios in the mid-latitudes may be approximately 200 hPa. During MOHAVE-2009 this pressure level corresponded roughly with the 12 km altitude level. The mean mixing ratio measured by frost point hygrometer at 12 km during MOHAVE-2009 was approximately 32 ppmv. Recall also Eq. (10), which indicates that a signal-dependent bias in Raman water vapor lidar measurements manifests itself as nearly a constant bias in mixing ratio space. Taking the 2σ value of 1.3 ppmv from Fig. 4 as the uncertainty in the mean value of the lower stratospheric mixing ratio of water vapor during the MOHAVE-2009 campaign and taking the value of 0.25 ppmv as the uncertainty in the determination of the bias correction, ζ_1 , and considering as well the uncertainty of the mean MLS value shown in Fig. 4 to be approximately 0.5 ppmv, the

propagated uncertainty of the correction becomes approximately 1.4 ppmv. Under the conditions present during the MOHAVE-2009 campaign, therefore, the use of MLS climatology to correct Raman water vapor lidar bias would contribute an additional uncertainty of approximately 4 % to the Raman lidar measurement of water vapor at the 200 hPa level in the upper troposphere, a region of the atmosphere where recent research (Whiteman et al., 2011b) indicates that random uncertainties of 50 % and more are acceptable for trend detection purposes.

8 Discussion

One of the goals of the NDACC Raman water vapor lidar effort is that of trend monitoring of water vapor in the atmosphere. Trends in lower stratospheric water vapor are clearly important in atmospheric science. For example, Solomon et al. (2010) recently showed that the 10 % decrease in stratospheric water vapor amounts that occurred after the year 2000 acted to slow the rate of increase in global surface temperature over the period 2000–2009 by about 25 %, compared to that which would have occurred due only to carbon dioxide and other greenhouse gases. However, as mentioned earlier, models predict that mean upper tropospheric water vapor mixing ratios are predicted to increase by up to 100 % or more over the coming century due to surface temperature increases, while mean lower stratospheric mixing ratio increases are anticipated to increase by less than half that amount (Soden et al., 2005; Boers and Meijgaard, 2009; Whiteman et al., 2011b). Trend detection in the upper troposphere is much more tolerant of additional noise in the time series than in the lower stratosphere (Whiteman et al., 2011b), and it needs to be noted, furthermore, that the random errors in Raman lidar data in general increase

rapidly, progressing from the upper troposphere into the lower stratosphere.

There are additional factors that are of concern when considering the possibility of developing a climate quality data record of lower stratospheric water vapor from Raman lidar. NDACC Raman lidars typically operate open to the atmosphere. The possibility of contamination of optics by fluorescence inducing-insects (Byrdegaard et al., 2009) or pollen (Saito et al., 2010), for example, or the possibility that airborne fluorescing particles (Gelbwachs and Birnbaum, 1973; Pan et al., 1999; Pinnick et al., 2004; Immler et al., 2005) or degraded hardware (Piironen, 1994) are biasing the data must be considered in developing procedures for data quality control. As shown here, these biases can be significant with respect to the mean LS water vapor mixing ratio. The possibility for system dependent biases to be present within the NDACC Raman lidar network is increased by the fact that each instrument within the network is unique both in terms of its hardware and software.

Therefore, considering the low signal-to-noise of the Raman lidar measurements in the lower stratosphere, the documented tendency for various factors to lead to generally wet biases in the measurements, and the need for higher precision measurements for trend detection in the lower stratosphere than the upper troposphere (Whiteman et al., 2011b), it seems more practical to target upper tropospheric water vapor as the prime measurement goal for Raman lidars. This can be aided by putting in place quality control procedures that can be used to check for the presence of biases in Raman lidar water vapor data on a regular basis and correct them.

With optimization, in particular through decreasing the noise due to skylight and from the detector itself, the measurements here and elsewhere (Whiteman et al., 2010; Leblanc et al., 2012a) demonstrate that Raman water vapor lidars can have significant sensitivity to water vapor in the lower stratosphere, although those measurements may be influenced by the biases mentioned earlier (cf. Sect. 4). The above details create the context for the proposed quality control procedure based on lower stratospheric water vapor climatology that may also be used for correction of data that possess signal-dependent bias. Given that biases could develop at any time due, for example, to deposition of material or degradation of system hardware, ongoing tests for the presence of biases should be performed. The procedure outlined here can be implemented without the need for a validation campaign involving FP launches and could be implemented routinely by Raman water vapor lidar systems within NDACC. This procedure could have been helpful during the full series of MOHAVE experiments that occurred during 2006, 2007 and 2009, during which wet biases in Raman water vapor lidars was one area of concentrated study.

As mentioned before, recent work indicates that the tolerance for water vapor measurement uncertainty is rather high when considering the task of trend detection in the upper troposphere. But it is important for all components of

the uncertainty budget to randomize over time (Whiteman et al., 2011b). The correction procedure described here, if implemented regularly as a part of the data processing protocol, would accomplish that for signal-dependent biases in the water vapor lidar measurements and thus help to normalize the upper tropospheric measurements across the NDACC network.

The need for data quality control in climate data records is illustrated by the efforts relating to quantifying atmospheric temperature trends using radiosondes. The historical record of radiosonde temperatures is plagued by various data quality issues (Seidel et al., 2004). A workshop convened to compare different data harmonization techniques concluded that the various techniques yielded sufficiently dissimilar results that no conclusions could be drawn as to the best method for adjusting the time series (Free et al., 2002). The point to be taken from this history is that, despite best intentions, errors do arise in data series that are intended to last for a period of decades and that data quality control procedures should be implemented at the initiation of a data record intended for climate monitoring purposes.

The use of mean lower stratospheric water vapor climatology as quality control and as a source of “tie point” for correcting upper tropospheric water vapor measurements over an extended period of time relies on future knowledge of LS water vapor mixing ratios. These values are anticipated to increase from current mid-latitude values of approximately 3–6 ppmv to perhaps 5–9 ppmv over the coming century (Oman et al., 2008; Whiteman et al., 2011b), while the anticipated changes in UT water vapor mixing ratio at the 200 hPa level may increase from the ~ 30 ppmv measured during MOHAVE-2009 to 60 ppmv or greater over the same period of time. Since future UT water vapor mixing ratios are expected to increase more than those in the LS, it is possible that the added uncertainty of using LS water vapor mixing ratio as a “tie point” for data correction will contribute a smaller amount to the uncertainty budget in the future than was the case here, where it amounted to an estimated 4 % at 200 hPa. Therefore, the future uncertainty of LS water vapor mixing ratios is not anticipated to diminish the utility of using mean LS values for quality controlling and correcting future Raman lidar upper tropospheric data.

9 Summary and conclusions

The participation of the mobile system known as ALVICE in the MOHAVE-2009 field campaign has been described. The ALVICE system was deployed to Table Mountain, CA, for MOHAVE-2009 with a large suite of remote sensing and in-situ instrumentation for quantifying water vapor and other atmospheric constituents. The rationale for, and development of, a correction technique for signal-dependent biases in Raman lidar water vapor measurements were described. Several topics pertinent to the calibration/validation

goals of MOHAVE-2009 campaign were presented in appendices. These appendices included information on the performance of the corrected Vaisala RS92 radiosonde during MOHAVE-2009; the development of a new radiosonde calibration technique that reduces the influence of atmospheric variability on the derived lidar calibration constant; the quantification of the full uncertainty budget of Raman lidar water vapor mixing ratio measurements; the development of specific data products for optimum comparison with radiosondes, satellites or other lidar systems; and a comparison of Integrated Precipitable Water (IPW) from lidar and GPS.

The research indicates that the ALVICE Raman lidar arrived for the MOHAVE-2009 field campaign free of any significant fluorescence. However, insects contaminated the ALVICE Raman lidar receiver early in the MOHAVE-2009 campaign and were the suspected cause of fluorescence in the lidar receiver that resulted in a wet bias in the ALVICE upper level measurements. Following this event, an optical window was installed over the lidar telescope, a large capacity fan was used to blow across the window during measurement periods and the window was washed daily. Given that analysis of the data indicates that the fluorescence contamination was essentially constant during the remainder of MOHAVE, it is possible that these measures are sufficient to prevent contamination of Raman water vapor lidar systems by fluorescing airborne material.

Examination of both MOHAVE data and those of other Raman lidar systems indicates that wet biases in upper tropospheric Raman lidar water vapor measurements are rather common. A mathematical model describing the physical process of signal-dependent bias was derived and applied to the ALVICE data to correct for the wet bias observed during MOHAVE-2009. Applying the wet bias correction derived here, using either actual measurements by FP or lower stratospheric climatology, resulted in corrected ALVICE profiles that agreed in the mean with FP to within 10% from 10–20 km.

The MOHAVE and WAVES experiments have shown that the elimination of all significant fluorescence from a Raman water vapor lidar system is possible through careful experimentation. However, each NDACC Raman lidar system is unique and uses custom-developed algorithms and therefore it should be expected that new challenges will emerge when the data from these systems receive careful inspection. How much effort will be required at these other sites to fully eliminate any wet bias that might be found? What is the likelihood that problems such as contamination by insects or pollen or degradation of hardware might create a wet bias following a cal/val campaign, as may have happened with the DOE/ARM lidar in the period 1998–2003? What are the chances that fluorescence of airborne particles (Gelbwachs and Birnbaum, 1973; Pan et al., 1999; Pinnick et al., 2004; Immel et al., 2005) or signal-induced noise (Piironen, 1994) might contaminate the data in ways that may not be correctable?

Consideration of these questions and the lessons from the difficulty in harmonizing temperature data records (Free et al., 2002) indicates that data quality controls should be developed within the NDACC Raman water vapor lidar effort now and put in place before data are archived for scientific use. Furthermore, archiving the complete raw data record should be considered a requirement since it would permit reprocessing in the future if problems are found to have developed in the time series.

We have demonstrated a potential quality control procedure for Raman water vapor lidar measurements of UTLS water vapor. This data quality check could be used to reject measurements that fall outside of reasonable limits, but given the many years of biased Raman lidar data that already exist and which could potentially be made useful for atmospheric studies in the upper troposphere, it makes more sense to correct the data. Correcting the data is also consistent with the recommendations of the international metrology community (JCGM/GUM, 2008). For the situation where the observed bias manifests itself as an essentially constant offset in ppmv space, the correction developed here for signal-dependent bias may be implemented. The use of this correction is shown to have a small effect on the uncertainty budget of upper tropospheric water vapor mixing ratio measurements during MOHAVE-2009. If biases are believed to be due to fluorescence of airborne particles (Pan et al., 1999; Pinnick et al., 2004; Immel et al., 2005), or other effects such as afterpulsing (Piironen, 1994), the data quality control aspect of the technique described here may be used to reject the data.

Raman water vapor lidar has proven to be an extremely valuable research tool for many scientific applications and it is expected that it can be so within the context of the NDACC network as well. However, experience has taught that measurements of water vapor with Raman lidar, particularly in the dry upper troposphere and lower stratosphere, present extreme challenges that must be addressed with great care. It is for this reason that we suggest that protocols be developed within the NDACC Raman water vapor lidar effort with the focus being on delivering quality controlled data products with sufficient accuracy for specific scientific applications.

Appendix A

RS92 RH accuracy and corrections

The accuracy of RS92 relative humidity (RH) measurements during MOHAVE-2009 is evaluated by comparing to two reference sensors of known accuracy: dual soundings with frost point hygrometers, and comparisons before launch with the Temperature-Humidity-Pressure reference system (THPref). The THPref is a surface reference system that provides NIST traceable measurements of temperature, relative humidity and pressure and contains a ventilated chamber for characterizing radiosonde accuracy prior to launch. The uncertainty

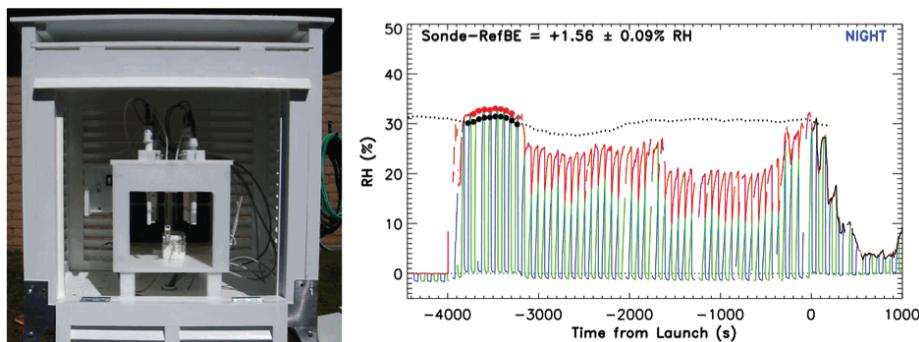


Fig. A1. On the left an RS92 is shown in the THPref, although normally the door is closed and faces north, and it is located away from buildings and other sources of heat or moisture. On the right is an example data comparison showing RH time series from the THPref (black dashed line) and RS92 (red prior to launch at $t = 0$, and black after launch). The large dots are 1-min averages during the comparison period when the RS92 is in the THPref, in this case indicating an RS92 mean bias relative to THPref of $+1.6\% \text{ RH}$. Green and blue curves represent the individual RS92 RH sensors that are alternately heated, while the other sensor measures where the RS92 RH data (red curve) is given by the combined measurement portions from each sensor. Note that prior to launch, recovery from a heating cycle is incomplete when the RS92 is not being ventilated in the THPref.

of the “best estimate” (averaged) THPref measurements are $\pm 0.1^\circ\text{C}$, $\pm 0.5\% \text{ RH}$, $\pm 0.08 \text{ hPa}$, respectively. The THPref (Fig. A1, left side) consists of six calibrated temperature and relative humidity probes in a fan-ventilated chamber within a naturally-ventilated instrument shelter, into which radiosondes are placed for comparative measurements prior to launch. The purpose of using multiple probes is both to reduce the random error of the measurement and to provide redundancy. The Reference system that travels with the ALVICE trailer was developed by Milo Scientific of Lafayette, CO, and is based on the original system developed by the DOE/ARM program and described in Miloshevich et al. (2009). We have augmented the instrument, which is now referred to as THPref, by addition of a precision barometer outside of the ventilated chamber. These pressure measurements are now coupled into the instrument data stream. During MOHAVE-2009 campaign, however, the P data were provided by the pressure sensor used with the SuomiNet GPS system. A comparison of the THPref RH measurements and those of a dual sensor RS92 radiosonde during the time of insertion of the sonde in the ventilated chamber is shown in the right panel of Fig. A1. Analysis of the THPref and RS92 raw pre-launch data gives the calibration bias of RS92 T and RH measurements relative to THPref under surface conditions. The THPref comparisons describe the RS92 accuracy under ambient conditions at the surface, and the comparisons with FPs and redundant RS92 sondes launched on the same balloons launched during MOHAVE-2009 (Hurst et al., 2011b) are used to characterize the RS92 accuracy in flight. During MOHAVE-2009 two versions of frost point hygrometers were launched, the CFH by TMF personnel and the NOAA FPH by NOAA personnel (Hurst et al., 2011b). On all of these launches an RS92 was flown as part of the payload. In the analysis of the RS92 RH measurement accuracy, the two versions of frost point

hygrometer were used equivalently. The frost point hygrometer data are henceforth referred to generically as FP data.

Corrections for several known sources of measurement error were applied to the RS92 RH data following a modified approach to that described by Miloshevich et al. (2009). Recent work (Kottayil et al., 2012) has demonstrated that generation of the empirical correction as a function of temperature, as opposed to pressure as was done previously, is more consistent with the calibration of the radiosonde sensors and improves performance in the tropical UTLS. It is this more recent version of the correction that is used here. A component of the correction is for sensor time-lag error and helps to recover vertical structure in the profile that is “smoothed” by slow sensor response at temperatures below about -45°C (Miloshevich et al., 2004). A correction is also applied for mean calibration bias, which is a moist bias in the lower troposphere and a dry bias at higher levels. A correction was also applied to the few daytime soundings for solar radiation dry bias caused by solar heating of the RH sensor. The bias corrections were derived to remove the mean bias of 2006–2007 vintage RS92 measurements relative to 3 reference sensors: frost point hygrometer above the 700 mb level, a THref system at the surface, and microwave radiometer (MWR) precipitable water measurements that represent mainly the lower troposphere. The RS92 RH accuracy and the correction for mean calibration bias are evaluated here by similar comparisons of the nighttime MOHAVE-2009 RS92 measurements to the FP and THPref reference sensors.

For 41 RS92-THPref comparisons during MOHAVE-2009, the mean and standard deviation of the RS92 RH and T biases were $+1.6 \pm 0.4\% \text{ RH}$ and $+0.09 \pm 0.16^\circ\text{C}$. In terms of water vapor mixing ratio (Fig. A2, left), the RS92 mean moist bias relative to THPref was $+6.2 \pm 3.4\%$, and it varies with RH from about $+12\%$ at $15\% \text{ RH}$ to $< 3\%$ above $50\% \text{ RH}$. The $\pm 0.5\% \text{ RH}$ uncertainty of the THPref

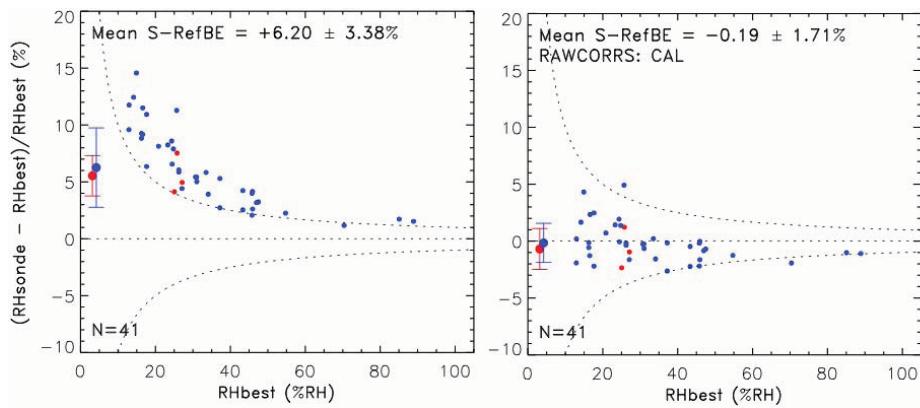


Fig. A2. Percentage difference between RS92 RH measurements and the mean of the 6 THPRef probes, shown as a function of RH for the original RS92 measurements (left panel) and after applying the correction for mean calibration bias (right panel). Blue indicates nighttime measurements and red indicates daytime. The curved dashed lines represent a difference of $\pm 1\%$ RH.

measurements is half the magnitude of the curved dashed lines. The correction for mean calibration bias (Fig. A2, right) mostly removes the RS92 mean bias and its RH-dependence, but it appears to overcorrect by about 1% for conditions above 30% RH. Most likely the over-correction reflects periodic re-calibration of the Vaisala factory references and possibly changes to the Vaisala calibration function since the 2006–2007 timeframe. This illustrates an important point that the mean calibration bias is expected to change with time and therefore a correction becomes out-of-date, unlike corrections for time-lag and solar radiation errors that only change with physical changes to the sensor or manufacturing process, or if Vaisala institutes their own corrections, which they have done for time-lag and solar radiation errors beginning with Digidata software version 3.64 released in December 2010.

For the 24 nighttime FP and RS92 dual soundings conducted during MOHAVE-2009, the RS92 had a moist mean bias relative to FP of about 5% in the lower troposphere, and at higher levels it had a dry mean bias that increased with height to a maximum of 40% in the tropopause region (Fig. A3, left panel). The calibration correction (Fig. A3, right panel) over-corrects in the lower troposphere by about the same magnitude as the measured mean bias, so it neither increases nor decreases the RS92 accuracy. At higher levels the calibration correction under-corrects by about 10% up to 12 km altitude, increasing to a maximum of 25% in the tropopause region. Time-lag error is also a factor in the UTLS when humidity gradients are steep, most notably around 19 km where the RH consistently decreases to low stratospheric values due to the increase in temperatures coupled with a nearly constant water vapor mixing ratio. Again it is thought that the calibration correction is less effective for 2009 radiosondes than for 2006–2007 radiosondes due to expected changes in the mean calibration bias with time.

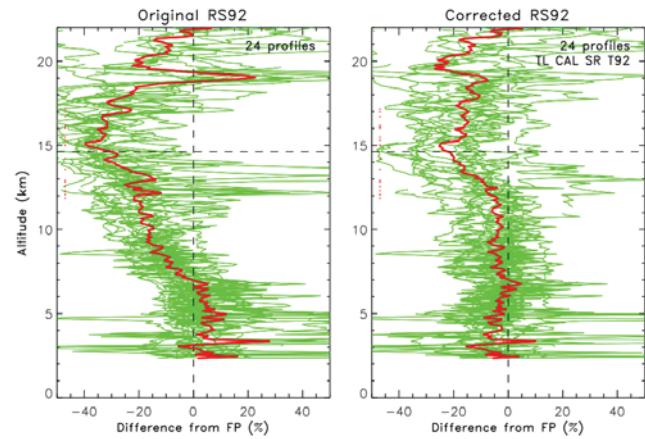


Fig. A3. Altitude profiles of the relative percentage difference between RS92 and frost point hygrometer (green) and the mean of all profiles (red) for the nighttime MOHAVE 2009 dual soundings, shown for the original RS92 measurements (left panel) and after applying the calibration and time-lag corrections (right panel). Dashed line is the mean tropopause height, and tiny red dots are the individual tropopause estimates.

Altitude profiles of the RS92 accuracy relative to FP can be misleading because they describe all RH conditions combined, whereas the RS92 calibration bias varies with RH. Extremely dry conditions were frequently encountered during MOHAVE-2009, when a small bias of 0.5% RH is relatively large (e.g., 50% bias at 1% RH). Figure A4 shows the same data as Fig. A3, but the RS92 bias relative to FP is shown in 4 RH intervals. The over-correction in the lower troposphere (rightmost dots) applies to all RH conditions, but the under-correction at lower pressures is only seen for conditions of $RH < 20\%$ and especially $RH < 10\%$.

The uncertainty in corrected RS92 RH measurements can be estimated from the bias uncertainty of $\pm (4\% + 0.5\% RH)$ given by Miloshevich et al. (2009), which is based on the

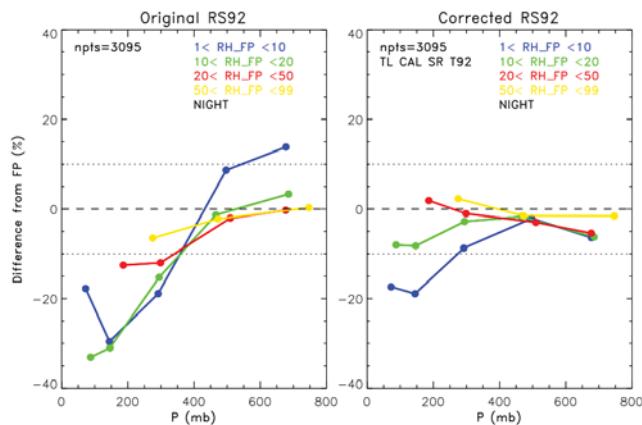


Fig. A4. RS92 mean bias relative to frost point hygrometer for the MOHAVE 2009 nighttime dual soundings, shown as a function of pressure (P) in 4 RH intervals, for the original RS92 measurements (left panel) and after applying the calibration and time-lag corrections (right panel).

size of the dataset used to derive the correction and the uncertainty in FP measurements. The constant 0.5 % RH component of uncertainty reflects the accuracy of the 0 % RH calibration and uncertainty in the Vaisala ground-check at 0 % RH. The random component of uncertainty (the 2σ sonde-to-sonde “production variability”) was estimated from dual RS92 soundings to be $\pm 3\%$ (for $\text{RH} > 10\%$) or $\pm 6\%$ (for $\text{RH} \leq 10\%$). The estimated total uncertainty for corrected RS92 measurements is the RMS sum of the bias and random components, or $\pm (5\% + 0.5\% \text{RH})$ for $\text{RH} > 10\%$ and $\pm (7\% + 0.5\% \text{RH})$ for $\text{RH} \leq 10\%$, which corresponds to an uncertainty of $\pm 6\%$ at 50 % RH, $\pm 10\%$ at 10 % RH, and $\pm 24\%$ at 3 % RH. The comparison to FP in Fig. A4 is consistent with this uncertainty estimate, but still there is evidence that the mean calibration bias for 2009 radiosondes has changed relative to the 2005–2008 radiosondes described by the above uncertainty estimate due to expected changes in the RS92 calibration with time, and indicating that the uncertainty estimate for 2005–2008 radiosondes is conservative.

A1 ALVICE lidar calibration, uncertainty budget, data products and IPW comparisons

Various ancillary material is presented in the following sections. A new technique for calibrating Raman lidar water vapor profiles with respect to radiosonde data was developed during the analysis of the MOHAVE-2009 data and is described here. For the first time, the full uncertainty budget of Raman water vapor lidar is estimated, including all significant sources of random and systematic error. Different versions of the data products were created to optimize comparisons with different types of sensors. The IPW of the fully corrected ALVICE data is compared with GPS IPW estimates. These are described in separate sections below.

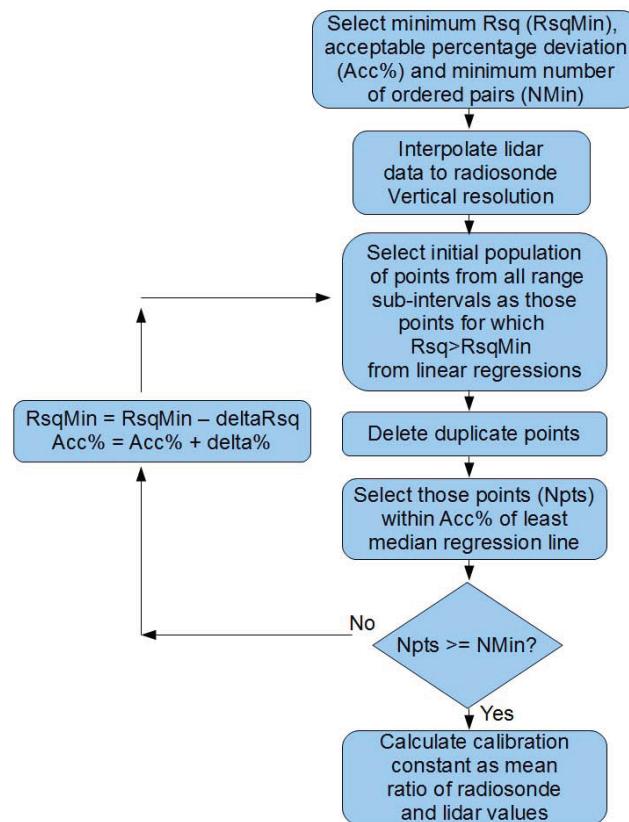


Fig. A5. Flow chart for the adaptive algorithm used to select geometrically similar portions of the lidar and radiosonde profiles for determining the lidar calibration constant. See text for more details. In the flow chart, “Rsq” refers to R^2 , and “deltaRsq” to ΔR^2 .

A2 Radiosonde based calibration technique

A traditional and very common method for calibrating a Raman water vapor lidar is to derive a calibration by comparison with a balloon borne radiosonde. The fact that the sonde does not sample the same atmosphere as the lidar coupled with the generally high variability of lower tropospheric water vapor, particularly at a mountain-top location like Table Mountain, raises concerns for how to best implement this technique. This question was discussed at the NDACC Raman water vapor lidar calibration workshop held in Greenbelt, MD, in May 2010. The algorithm used here for lidar calibration with respect to radiosonde profiles was developed as an outgrowth of the concerns expressed at the workshop about the influence of atmospheric variability on calibration of Raman lidar with respect to radiosonde. The algorithm presented here attempts to account for this influence on a profile-by-profile basis. A flow chart of the algorithm is given in Fig. A5 and next we provide a description of how the algorithm operates. Supplemental material is available that contains a fully operational example of the routine that is implemented in *Mathematica*.

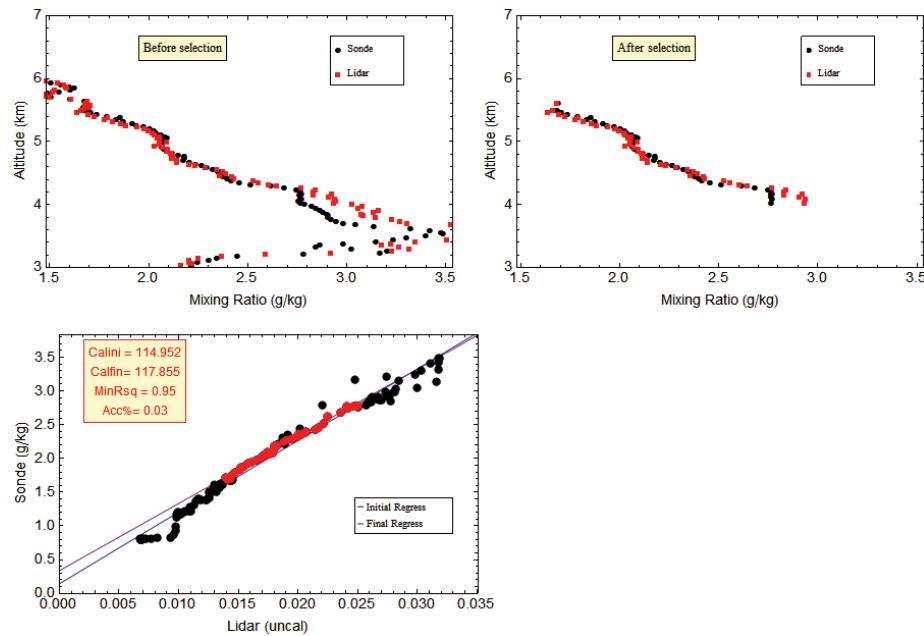


Fig. A6. Demonstration of the adaptive radiosonde calibration routine. In the upper left is shown both lidar (red) and radiosonde (black) profiles without data filtering. In the upper right is shown the sets of ordered pairs that are selected by the algorithm described in the text. On the bottom is shown the regression lines of the original set of lidar and sonde ordered pairs in black and the finally selected set in red.

The algorithm works on the assumption that if the lidar and the radiosonde are sampling the same atmosphere within a given range cell, the shape of the mixing ratio curves in that portion of their respective profiles will be geometrically similar. The mean proportionality constant between those geometrically similar curves will be the calibration constant derived from that portion of the profile. The algorithm proceeds over a specified range of minimum and maximum altitudes and performs a series of linear least squares and least median of squares regressions for subsets of the data within these range limits. These regressions are used to gauge the similarity of the shapes of the curves through the R^2 values of the linear regressions and to eliminate outliers through the use of the least median regressions. As implemented here, ordered pairs of lidar and radiosonde data from 0.5 km range segments are regressed within the height range of 3 and 7 km m.s.l. Ordered pairs are accepted as members of the final set of data used to determine the calibration value if, first, they were part of a regression with sufficiently high R^2 , and second, if an individual ordered pair is within a certain percentage of a least median of squares fit line. The algorithm is adaptive so that, if after the first pass through the profile an insufficient number of points has qualified for the calibration determination, the acceptance criteria for R^2 and maximum deviation from the least median of squares prediction line are relaxed and the process is repeated until the desired number of ordered pairs is obtained. The final calibration value is determined by the mean ratio of the lidar and radiosonde data and not by the slope of a regression line.

An example of the results of this routine is shown in Fig. A6 using a radiosonde comparison on the night of 21 October 2009. The plot in the upper left shows the radiosonde and lidar profiles in the region of interest. The calibration constant derived by the algorithm for this comparison has been used here for scaling the lidar data. In the upper right are plotted just the ordered pairs that were finally accepted by the algorithm for use in determining the calibration constant. Here the ability to select the parts of the profile that show similar geometrical shapes is demonstrated. In the lower left is shown a comparison of the linear regression of all points within the 3 to 7 km range (in black) along with only the qualifying points used in the calibration determination. In this plot also can be found the calibration constants calculated with both the full population of points before any selection (yielding a value of 114.92 g kg^{-1}) and just those that qualified after selection (117.86 g kg^{-1}). For these examples, the minimum number of order pairs required to perform the calibration was set at 30. This corresponds to a range in the atmosphere of 900 m.

A study of the calibration constants derived using this technique indicated that the standard deviation of the constants decreased as the minimum acceptable R^2 value required to obtain 30 ordered pairs increased. This implies that the more geometrically similar the lidar and radiosonde profiles are, the more stable the derived calibration constant. Based on 33 comparisons of lidar and radiosonde profiles, it was also found that restricting the RH values from radiosonde to values above 5 % RH further decreased the

standard deviation of the derived constants from 9.3 to 8.5 %. Low RH conditions are known to cause the absolute errors in Vaisala RS92 RH data to increase (Miloshevich et al., 2009) and are best avoided for use in lidar calibration. Calibration constants were determined separately using RS92 and CFH data. However, there are indications that the corrections applied to the Vaisala RS92 data may have resulted in an over-correction during MOHAVE-2009 (cf. Appendix A). Also, the CFH can possess a wet bias in the lower troposphere (Miloshevich et al., 2009). Therefore, the final calibration value used for the processing of the ALVICE Raman lidar data was determined by averaging the calibration constants determined from corrected RS92 and frost point hygrometer (FP).

Prior to data runs on 10 nights during MOHAVE-2009, a stabilized tungsten lamp was used to directly illuminate the lidar receiver in an implementation of the “hybrid technique” (Leblanc and McDermid, 2008, 2011; Whiteman et al., 2011a). The ratio of the signals in the water vapor and nitrogen channels due to the lamp illumination is recorded with the goal of using this ratio to stabilize the calibration obtained from radiosonde. The standard deviation of radiosonde-derived calibration constants was 5.5 % without implementing the hybrid technique and was 5.6 % by implementing the hybrid technique. Thus, no improvement in calibration stability was achieved by use of the recorded lamp ratios. The lamp ratios recorded during the first 5 nights of operation showed a standard deviation of approximately 5 %, while the lamp ratios for the remaining 5 nights showed a standard deviation of less than 1 %. Comparison of ALVICE water vapor profiles with those of the other lidars and balloon-borne sensors do not indicate that the calibration constant of the lidar was truly more variable during the early part of the mission versus the latter. Therefore, we speculate that the considerably higher standard deviation during the first 5 nights may have been due to slight shifting of position of some of the residual insect material at the location on the telescope primary mirror that was being sampled by a single, fixed calibration lamp implemented during MOHAVE-2009 (Whiteman et al., 2011a). Since that time, the ability to scan the full telescope aperture with the calibration lamp has been added to the ALVICE system, supporting an independent calibration (Venable et al., 2011) of the water vapor products.

A3 ALVICE Raman lidar water vapor mixing ratio uncertainty budget

The Guide to the Expression of Uncertainty in Measurements (JCGM/GUM, 2008) states that “In practice, the required specification or definition of the measurand is dictated by the required accuracy of measurement. The measurand should be defined with sufficient completeness with respect to the required accuracy so that for all practical purposes associated with the measurement its value is unique.” To our knowledge, however, there have been no previous efforts to quantify the

total uncertainty of Raman lidar water vapor mixing ratio measurements including the effects of random error and all significant sources of systematic error. Given that there are various studies in which the water vapor profile data supplied here may be used (trend detection in the UT, trend detection in the LS, satellite retrieval validation, satellite radiance validation), we therefore attempt to provide in these datafiles an estimate of the total uncertainty due to all significant influence factors. The usefulness of an individual profile is then determined by considering the total error of the profiles as a function of altitude and the tolerance for error of the specific scientific study desired. Furthermore, a total uncertainty budget is needed for the purposes of data assimilation and in order to assess the influence of data error on time to detect trends (Whiteman et al., 2011b).

Therefore, for all versions of the ALVICE data analyzed for the MOHAVE campaign, the total uncertainty is estimated by including the following contributions. All uncertainties are systematic except for #1 below. The total propagated uncertainty was determined assuming all errors are uncorrelated.

1. Uncertainty due to random error in the lidar data calculated assuming Poisson statistics.
2. Uncertainty in the correction for the lidar overlap function, which is as large as 10 % in the lowest levels of the atmosphere.
3. Uncertainty of the calibration source. Based on absolute accuracy information found in Miloshevich et al. (2009) and Vömel et al. (2007b), the accuracy of the calibration sources varies between 4 % in the lower troposphere to 10 % in the lower stratosphere.
4. Uncertainty in the transfer of the calibration source to the lidar profile, which is estimated at 1–2 %.
5. Uncertainty in the correction for the temperature-dependence of Raman scattering, which is estimated to be less than 1 %.
6. Uncertainty in the correction for differential transmission of the atmosphere at the two Raman wavelengths, which is estimated to be less than 1 %.
7. Uncertainty in the correction for fluorescence. The correction for fluorescence is determined in Sect. 6 to be essentially constant during the campaign. To be conservative, an uncertainty of 0.25 ppmv was used in the data files.

We note that the uncertainty due to correction for photon pileup, i.e., the likelihood that two photons arrive spaced closely enough in time to be counted as a single event (Donovan et al., 1993; Whiteman et al., 1992), is excluded due to the use of a combination of analog-to-digital and photon counting data acquisition. This permits limiting the count

rates to less than 10 MHz and also provides direct information on how to correct the photon counting data (Whiteman et al., 2006a; Newsom et al., 2009). The uncertainty in the correction is therefore much less than 1 % and is excluded from the list above.

A4 Data products tailored for different applications

Different versions of the mixing ratio profiles are created to aid comparisons with different sensors. Simple averaging schemes are used to aid comparison of lidar profiles with other lidar instruments, while more sophisticated temporal and spatial smoothing is used to optimize comparison of lidar data with radiosonde or satellite. All profiles except the raw data profile have been smoothed with a moving window filter the size of which varies from 30 m in the lowest part of the atmosphere to a maximum of 1200 m for altitudes beyond 12 km. For altitudes above 12 km, the ALVICE lidar profiles possess vertical resolution of approximately 0.5–0.75 km using the Verein Deutscher Ingeneure (VDI4210, 1999) definition of vertical resolution for lidar systems. There also are four corrections that are applied to all water vapor profiles in these data files and the values of these corrections are given individually in the data files. The corrections are for water vapor mixing ratio overlap-dependence, temperature-dependence of Raman scattering, atmospheric differential transmission and the signal-dependent correction that is described in Sect. 5. The different vertically smoothed data products contained in the data files and their optimized uses are

1. a one-hour summation beginning at the time of the radiosonde launch, independent of altitude – convenient for comparison with another lidar instrument.
2. a summation of all available data for a given night – convenient for comparison with another lidar instrument.
3. variable temporal and spatial smoothing to limit the random error, if possible, to less than 5 % – best for capturing short term variability in the water vapor field.
4. “best estimate” product, which combines the above three processing schemes. The first three processing methods listed above have been studied previously (Whiteman et al., 2006c) and method 3 has proven to be preferred for satellite validation because of its ability to capture the high variability of lower tropospheric water vapor. Therefore, for MOHAVE-2009, this 4th data product, referred to as the “best estimate” product, was created. Borrowing from the philosophy of the DOE/ARM program (Ackerman and Stokes, 2003; Tobin et al., 2006), whereby the best available data are supplied as a function of altitude, the ALVICE best estimate profile is comprised of the 3 profiles listed above, including all corrections, along with a surface reference point derived from the surface reference data shown in

Fig. 1. This best estimate product is intended to be the most accurate profile of atmospheric water vapor at the time of the radiosonde launch from the ALVICE lidar perspective.

An illustration of the use of the best estimate product is provided in Fig. A7. In the left panel is shown a single profile comparison of the ALVICE best estimate profile, 1 h sum profile and all-night profiles for the radiosonde launch on 25 October at 03:55 UT. There was a large gradient in water vapor near the surface that filled in as time passed so that the best estimate product, which used as little as two minutes of data in the composite profile in the lower atmosphere, was able to capture this dry feature. The 1 h and all-night sums show large differences from the sonde. The mean bias and RMS for 33 Vaisala RS92 and ALVICE best estimate, 1 h and all-night profile comparisons are shown in the middle and on the right of Fig. A7. The RMS difference between sonde and best estimate product is consistently less than the 1 h and all-night averages up to an altitude of approximately 7 km. The mean RMS for the best estimate, 1 h and all-night averages up to an altitude of 7 km are, respectively, 17, 22 and 50 %. The mean biases for the three different ALVICE profile products are 1.3, 2.5 and 25 %, respectively. Thus, for the entire MOHAVE-2009 campaign both in terms of RMS and bias, the technique of using variable temporal averaging is found to agree better with radiosonde below 7 km. The small tendency for the 1 h data to show lower RMS and bias than the best estimate product above 6–7 km may be an indication that the averaging time used in the best estimate product in this altitude range should be increased. These results indicate that, in general, the best estimate profile provides a better representation of the state of the atmosphere at the time of the radiosonde launch than either the 1 h or all-night profiles. The large mean bias of the all-night data shown in the right panel of the figure implies that cases such as that illustrated in the left panel of the figure where a dry layer filled in during the night occurred with some frequency during MOHAVE-2009 and that care should be taken in comparing extended averages of lidar data with sensors that measure more rapidly such as those carried by radiosonde or satellite.

A5 Total column water comparisons using GPS, radiosonde and lidar

GPS total column water measurements can be considered a redundant source of calibration for Raman water vapor lidar, according to recommendations of the GCOS Reference Upper Air Network (GRUAN) operational protocols. Therefore, it is interesting to compare the IPW measurements of the GPS that accompanies ALVICE on deployments with the RS92 and the lidar IPW during the MOHAVE campaign to assess the possible use of each as a calibration source for the Raman water vapor lidar. A more detailed comparison of the total column water measurements made during MOHAVE-2009 is performed in Leblanc et al. (2011). Therefore, here

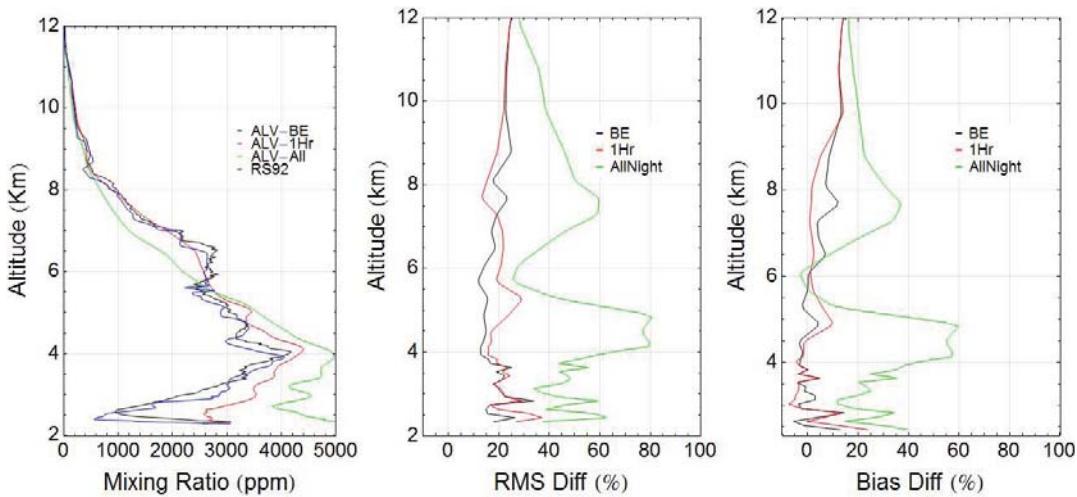


Fig. A7. Left panel: comparison of best estimate, 1 h sum and “all-night” ALVICE profiles and the RS92 radiosonde launch on 25 October at 03:55 UT. Mean RMS (middle panel) and bias (right panel) of the different ALVICE data products compared with all 33 RS92 radiosonde launches. The RMS and bias are calculated from the individual lidar-radiosonde comparisons with 100 m vertical binning up to 6 km and 400 m binning above 6 km.

we give just a brief description of the measurements made by the GPS system that accompanies the ALVICE trailer on its mobile deployments and compare those measurements with lidar and radiosonde. This system is the only “rover” within the international SuomiNet GPS network (Ware et al., 2000). It has been used previously as the calibration source for mobile Raman lidar water vapor measurements made during the International Water Vapor Project (IHOP) campaign held in the US Midwest in 2002 (Whiteman et al., 2006a,b) using a similar calibration procedure as that developed within the DOE/ARM program (Turner et al., 2002). During IHOP, comparisons made between lidar and frost point hygrometer (Whiteman et al., 2006a) and airborne water vapor lidars (Behrendt et al., 2007) showed agreement to generally better than 10 %.

During MOHAVE-2009, the SuomiNet GPS system with ID SA65 operated continuously from 10 October through most of 27 October 2009. The integrated precipitable water and pressure data from SA65 were combined with the temperature and RH data from the THPref instrument to provide a surface reference datafile containing RH, T , P , IPW and water vapor mixing ratio with a 5 min temporal resolution for the period 10–27 October, as previously shown in Fig. 1.

In general, Raman lidar systems have difficulty with measurements at close ranges due to the influence of the lidar system overlap function (Harms et al., 1978). This can be seen to varying degrees in the MOHAVE-2009 archived data from the different Raman lidars. A correction has been applied to the ALVICE Raman lidar data for these near field effects. By carefully selecting radiosonde profiles in a manner similar to that described in Sect. A2, an overlap correction was derived as the mean ratio of radiosonde and lidar data for the selected profiles. This technique has been used successfully

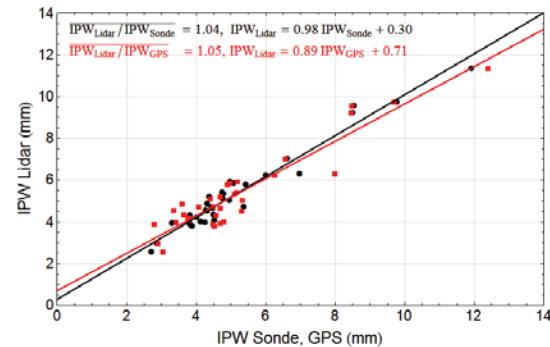


Fig. A8. Comparison of integrated precipitable water from the ALVICE lidar using the best estimate product and corrected RS92 (black) and SuomiNet GPS (red). The lidar results are 4–5 % moister than both the radiosonde and GPS.

previously (Turner et al., 2002; Whiteman et al., 2006c). The correction extends from the first reported lidar range bin, which is generally at 60 m a.g.l. to 1.2 km a.g.l. and reaches a maximum of ~40 % at the closest ranges. The uncertainty in the correction is estimated to be approximately 10 % based on the standard deviation of the lidar-radiosonde comparisons in the overlap region. The combination of overlap-corrected lidar mixing ratio extending down to 60 m, along with a ground point derived from the ALVICE surface reference data, improves the IPW calculated from lidar. It should be noted that as the alignment between the laser and the telescope field of view changes, the overlap function can be expected to change. The ALVICE Raman lidar, however, includes an automated bore-site alignment system that maintains the alignment to within < 20 microradians. Because of

this stability of alignment, we do not see significant changes in the overlap function due to alignment changes.

The IPW from SA65 and RS92 radiosonde were compared with the integrated precipitable water calculated from the best estimate ALVICE Raman lidar water vapor mixing ratio profile. The results are shown in Fig. A8. The mean ratio of IPW derived from lidar and IPW from RS92 and GPS were found to be 1.04 and 1.05, respectively, indicating that the lidar was approximately 5 % moister than either of these instruments. This moistness with respect to the corrected RS92 is consistent with the lower tropospheric dryness of the corrected RS92 versus FP shown in Fig. A3.

Supplementary material related to this article is available online at: <http://www.atmos-meas-tech.net/5/2893/2012/amt-5-2893-2012-supplement.zip>.

Acknowledgements. The authors wish to acknowledge the NASA Atmospheric Composition Program, managed by Ken Jucks, for support of these efforts. The mention of a particular vendor does not constitute an endorsement by NASA.

Edited by: G. Pappalardo

References

Ackerman, T. and Stokes, G.: The Atmospheric Radiation Measurement program, *Phys. Today*, 56, 38–45, 2003.

Behrendt, A. and Reichardt, J.: Atmospheric temperature profiling in the presence of clouds with a pure rotational Raman lidar by use of an interference-filter-based polychromator, *Appl. Optics*, 39, 1372–1378, 2000.

Behrendt, A., Wulfmeyer, V., Di Girolamo, P., Kiemle, C., Bauer, H.-S., Schaberl, T., Summa, D., Whiteman, D. N., Demoz, B. B., Browell, E. V., Ismail, S., Ferrare, R., Kooi, S., Ehret, G., and Wang, J.: Intercomparison of water vapor data measured with lidar during IHOP_2002, Part 1: Airborne to ground-based lidar systems and comparisons with chilled-mirror hygrometer radiosondes, *J. Atmos. Ocean. Tech.*, 24, 3–21, doi:10.1175/JTECH1924.1, 2007.

Boens, N., Qin, W., Basarić, N., Hofkens, J., Ameloot, M., Pouget, J., Lefèvre, J. P., Valeur, B., Gratton, E., vandeVen, M., Silva Jr., N. D., Engelborghs, Y., Willaert, K., Sillen, A., Rumbles, G., Phillips, D., Visser, A. J., van Hoek, A., Lakowicz, J. R., Malak, H., Gryczynski, I., Szabo, A. G., Krajcik, D. T., Tamai, N., and Miura, A.: Fluorescence lifetime standards for time and frequency domain fluorescence spectroscopy, *Anal. Chem.*, 79, 2137–2149, 2007.

Boers, R. and van Meijgaard, E.: What are the demands on an observational program to detect trends in upper tropospheric water vapor anticipated in the 21st century?, *Geophys. Res. Lett.*, 36, L19806, doi:10.1029/2009GL040044, 2009.

Brydegaard, M., Guan, Z., Wellenreuther, M., and Svanberg, S.: Insect monitoring with fluorescence lidar techniques: feasibility study, *Appl. Optics*, 48, 5668–5677, 2009.

Di Girolamo, P., Marchese, R., Whiteman, D. N., and Demoz, B. B.: Rotational Raman lidar measurements of atmospheric temperature in the UV, *Geophys. Res. Lett.*, 31, L01106, doi:10.1029/2003GL018342, 2004.

Dionisi, D., Congeduti, F., Liberti, G. L., and Cardillo, F.: Calibration of a Multichannel Water Vapor Raman Lidar through Noncollocated Operational Soundings: Optimization and Characterization of Accuracy and Variability, *J. Atmos. Ocean. Tech.*, 27, 108–121, doi:10.1175/2009JTECHA1327.1, 2009.

Donovan, D. P., Whiteway, J. A., and Carswell, A. I.: Correction for nonlinear photon-counting effects in lidar systems, *Appl. Optics*, 32, 6742–6753, 1993.

Dvorak, M. A., Oswald, G. A., Van Benthem, M. H., and Gillispie, G. D.: On-the-Fly Fluorescence Lifetime Determination with Total Emission Detection in HPLC, *Anal. Chem.*, 69, 3458–3464, 1997.

Ferrare, R. A., Browell, E. V., Ismail, S., Kooi, S. A., Brasseur, L. H., Brackett, V. G., Clayton, M. B., Barrick, J. D. W., Diskin, G. S., Goldsmith, J. E. M., Lesht, B. M., Podolske, J. R., Sachse, G. W., Schmidlin, F. J., Turner, D. D., Whiteman, D. N., Tobin, D., Miloshevich, L. M., Revercomb, H. E., Demoz, B. B., and Di Girolamo, P.: Characterization of Upper-Troposphere Water Vapor Measurements during AFWEX Using LASE, *J. Atmos. Ocean. Tech.*, 21, 1790–1808, 2004.

Free, M., Durre, I., Aguilar, E., Seidel, D., Peterson, T. C., EskrIDGE, R. E., Luers, J. K., Parker, D., Gordon, M., Lanzante, J., Klein, S., Christy, J., Schroeder, S., Soden, B., McMillin, L. M., and Weatherhead, E.: Creating Climate Reference Datasets – CARDS Workshop on Adjusting Radiosonde Temperature Data for Climate Monitoring, *B. Am. Meteorol. Soc.*, 891–899, 2002.

Gelbwachs, J. and Birnbaum, M.: Fluorescence of Atmospheric Aerosols and Lidar Implications, *Appl. Optics*, 12, 2442–2447, 1973.

Harms, J., Lahmann, W., and Weitkamp, C.: Geometrical Compression of Lidar Return Signals, *Appl. Optics*, 17, 1131–1135, 1978.

Held, I. M. and Soden, B. J.: Robust Responses of the Hydrological Cycle to Global Warming, *J. Climate*, 19, 5686–5699, 2006.

Hurst, D. F., Oltmans, S. J., Vömel, H., Rosenlof, K. H., Davis, S. M., Ray, E. A., Hall, E. G., and Jordan, A. F.: Stratospheric water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record, *J. Geophys. Res.*, 116, D02306, doi:10.1029/2010JD015065, 2011a.

Hurst, D. F., Hall, E. G., Jordan, A. F., Miloshevich, L. M., Whiteman, D. N., Leblanc, T., Walsh, D., Vömel, H., and Oltmans, S. J.: Comparisons of temperature, pressure and humidity measurements by balloon-borne radiosondes and frost point hygrometers during MOHAVE-2009, *Atmos. Meas. Tech.*, 4, 2777–2793, doi:10.5194/amt-4-2777-2011, 2011b.

Immler, F., Engelbart, D., and Schrems, O.: Fluorescence from atmospheric aerosol detected by a lidar indicates biogenic particles in the lowermost stratosphere, *Atmos. Chem. Phys.*, 5, 345–355, doi:10.5194/acp-5-345-2005, 2005.

JCGM/WG 1: Evaluation of measurement data – Guide to the expression of uncertainty in measurement, International Bureau of Weights and Measures/Bureau International des Poids et Mesures, Working Group 1 of the Joint Committee for Guides in Metrology, http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf (last access: November 2012), 2008.

Kottayil, A., Buehler, S. A., John, V. O., Miloshevich, L. M., Milz, M., and Holl, G.: On the Importance of Vaisala RS92 Radiosonde Humidity Corrections for a Better Agreement between Measured and Modeled Satellite Radiances, *J. Atmos. Ocean. Tech.*, 29, 248–259, doi:10.1175/JTECH-D-11-00080.1, 2012.

Lambert, A., Read, W. G., Livesey, N. J., Santee, M. L., Manney, G. L., Froidevaux, L., Wu, D. L., Schwartz, M. J., Pumphrey, M., Jimenez, C., Nedoluha, G. E., Cofield, R. E., Cuddy, D. T., Daffer, W. F., Drouin, B. J., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Pickett, H. M., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Jucks, K. W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., Murtagh, D., Elkins, J. W., and Atlas, E.: Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements, *J. Geophys. Res.*, 112, D24S36, doi:10.1029/2007JD008724, 2007.

Leblanc, T. and McDermid, I. S.: Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements, *Appl. Optics*, 47, 30, 5592–5603, 2008.

Leblanc, T. and McDermid, I. S.: Reply to “Comments on Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements by Whiteman et al.”, *Appl. Optics*, 50, 2177–2178, 2011.

Leblanc, T., Walsh, T. D., McDermid, I. S., Toon, G. C., Blavier, J.-F., Haines, B., Read, W. G., Herman, B., Fetzer, E., Sander, S., Pongetti, T., Whiteman, D. N., McGee, T. G., Twigg, L., Sunnicht, G., Venable, D., Calhoun, M., Dirisu, A., Hurst, D., Jordan, A., Hall, E., Miloshevich, L., Vömel, H., Straub, C., Kampfer, N., Nedoluha, G. E., Gomez, R. M., Holub, K., Gutman, S., Braun, J., Vanhove, T., Stiller, G., and Hauchecorne, A.: Measurements of Humidity in the Atmosphere and Validation Experiments (MOHAVE)-2009: overview of campaign operations and results, *Atmos. Meas. Tech.*, 4, 2579–2605, doi:10.5194/amt-4-2579-2011, 2011.

Leblanc, T., McDermid, I. S., and Walsh, T. D.: Ground-based water vapor raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring, *Atmos. Meas. Tech.*, 5, 17–36, doi:10.5194/amt-5-17-2012, 2012a.

Leblanc, T., McDermid, I. S., and Walsh, T. D.: Ground-based water vapor raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring, *Atmos. Meas. Tech.*, 5, 17–36, doi:10.5194/amt-5-17-2012, 2012b.

Livesey, N. J., Read, W. G., Froidevaux, L., Lambert, A., Manney, G. L., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Cofield, R. E., Cuddy, D. T., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Knosp, B. W., Stek, P. C., Wagner, P. A., and Wu, D. L.: Version 3.3 level 2 data quality and description document, available at: http://mls.jpl.nasa.gov/data/v3-3_data-quality_document.pdf, last access: November 2012, Jet Propulsion Laboratory/California Institute of Technology, JPL, 33509, 156 pp., 2011.

Mastenbrook, H. J. and Dinger, J. E.: Distribution of water vapor in the stratosphere, *J. Geophys. Res.*, 66, 1437–1444, 1961.

Miloshevich, L. M., Paukkunen, A., Vömel, H., and Oltmans, S. J.: Development and validation of a time-lag correction for Vaisala radiosonde humidity measurements, *J. Atmos. Ocean. Tech.*, 21, 1305–1327, 2004.

Miloshevich, L. M., Vömel, H., Whiteman, D. N., Lesht, B. M., Schmidlin, F. J., and Russo, F.: Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation, *J. Geophys. Res.*, 111, D09S10, doi:10.1029/2005JD006083, 2006.

Miloshevich, L. M., Vömel, H., Whiteman, D. N., and Leblanc, T.: Accuracy Assessment and Correction of Vaisala RS92 radiosonde water vapor measurements, *J. Geophys. Res.*, 114, D11305, doi:10.1029/2008JD011565, 2009.

Newsom, R. K., Turner, D. D., Mielke, B., Clayton, M. F., Ferrare, R., and Sivaraman, C.: Simultaneous analog and photon counting detection for Raman lidar, *Appl. Optics*, 48, 3903–3914, 2009.

Oman, L., Waugh, D., Pawson, S., Stolarski, R., and Nielsen, J.: Understanding the Changes of Stratospheric Water Vapor in Coupled Chemistry–Climate Model Simulations, *J. Atmos. Sci.*, 65, 3278–3291, 2008.

Pan, Y.-Le, Holler, S., Chang, R. K., Hill, S. C., Pinnick, R. G., Niles, S., and Bottiger, J. R.: Single-shot fluorescence spectra of individual micrometer-sized bioaerosols illuminated by a 351- or a 266-nm ultraviolet laser, *Opt. Lett.*, 24, 116–118, 1999.

Piironen, P.: Data System in “A High Spectral Resolution Lidar Based on an Iodine Absorption Filter”, Ph.D. Thesis, University of Joensuu, Department of Physics, Finland, 1994.

Pinnick, R. G., Hill, S. C., Pan, Y.-Le, and Chang, R. K.: Fluorescence spectra of atmospheric aerosol at Adelphi, Maryland, USA: measurement and classification of single particles containing organic carbon, *Atmos. Environ.*, 38, 1657–1672, 2004.

Sakai, T., Nagai, T., Nakazato, M., Matsumura, T., Orikasa, N., and Shoji, Y.: Comparisons of Raman lidar measurements of tropospheric water vapor profiles with radiosondes, hygrometers on meteorological observation tower, and GPS at Tsukuba, Japan, *J. Atmos. Ocean. Tech.*, 24, 1407–1423, 2007.

Saito, Y., Kobayashi, K., and Kobayashi, F.: Laser-induced Fluorescence Spectrum (LIFS) Lidar for Remote Detection of Biological Substances Surrounding the “LivingSphere”, The 8th International Symposium on Advanced Environmental Monitoring, Hokkaido, Japan, 2010.

Seidel, D. J., Angell, J. K., Christy, J., Free, M., Klein, S. A., Lanzante, J. R., Mears, C., Parker, D., Schable, M., Spencer, R., Stern, A., Thorne, P., and Wentz, F.: Uncertainty in signals of large-scale climate variations in radiosonde and satellite upper-air temperature datasets, *J. Climate*, 17, 2225–2240, 2004.

Sherlock, V., Garnier, A., Hauchecorne, A., and Keckhut, P.: Implementation and validation of a Raman lidar measurement of middle and upper tropospheric water vapor, *Appl. Optics*, 38, 5838–5850, 1999.

Soden, B. J., Jackson, D. L., Ramaswamy, V., Schwarzkopf, M. D., and Huang, X.: The radiative signature of upper tropospheric moistening, *Science*, 310, 841–844, 2005.

Solomon, S., Rosenlof, K., Portmann, R., Daniel, J., Davis, S., Sanford, T., and Plattner, G.-K.: Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, *Science*, 327, 1219–1223, 2010.

Stiller, G. P., Kiefer, M., Eckert, E., von Clarmann, T., Kellmann, S., García-Comas, M., Funke, B., Leblanc, T., Fetzer, E., Froidevaux, L., Gomez, M., Hall, E., Hurst, D., Jordan, A., Kämpfer, N., Lambert, A., McDermid, I. S., McGee, T., Miloshevich, L., Nedoluha, G., Read, W., Schneider, M., Schwartz, M., Straub, C., Toon, G., Twigg, L. W., Walker, K., and Whiteman, D. N.:

Validation of MIPAS IMK/IAA temperature, water vapor, and ozone profiles with MOHAVE-2009 campaign measurements, *Atmos. Meas. Tech.*, 5, 289–320, doi:10.5194/amt-5-289-2012, 2012.

Tobin, D. C., Revercomb, H. E., Knuteson, R. O., Lesht, B. M., Strow, L. L., Hannon, S. E., Feltz, W. F., Moy, L. A., Fetzer, E. J., and Cress, T. S.: Atmospheric Radiation Measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water vapor retrieval validation, *J. Geophys. Res.*, 111, D09S14, doi:10.1029/2005JD006103, 2006.

Turner, D. D., Ferrare, R. A., Heilman Breasseur, L. A., Feltz, W. F., and Tooman, T. P.: Automated Retrievals of Water Vapor and Aerosol Profiles from an Operational Raman Lidar, *J. Atmos. Ocean. Tech.*, 19, 37–50, 2000.

VDI4210: Blatt 1, Remote Sensing Atmospheric Measurements with LIDAR – Measuring gaseous air pollution with DAS Lidar, ICS 13.040.20, 17.180.01, Verein Deutscher Ingenieure, 1999.

Venable, D. D., Whiteman, D. N., Calhoun, M. N., Dirisu, A. O., Connell, R. M., and Landulfo, E.: A Lamp Mapping Technique for Independent Determination of the Water Vapor Mixing Ratio Calibration Factor for a Raman Lidar System, *Appl. Optics*, 50, 4622–4632, 2011.

Vömel, H., David, D. E., and Smith, K.: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations, *J. Geophys. Res.*, 112, D08305, doi:10.1029/2006JD007224, 2007a.

Vömel, H., Barnes, J. E., Forno, R. N., Fujiwara, M., Hasebe, F., Iwasaki, S., Kivi, R., Komala, N., Kyro, E., Leblanc, T., Morel, B., Ogino, S.-Y., Read, W. G., Ryan, S. C., Saraspriya, S., Selkirk, H., Shiotani, M., Valverde Canossa, J., and Whiteman, D. N.: Validation of Aura Microwave Limb Sounder water vapor by balloonborne Cryogenic Frost point Hygrometer measurements, *J. Geophys. Res.*, 112, D24S37, doi:10.1029/2007JD008698, 2007b.

Wang, J., Carlson, D. J., Parsons, D. B., Hock, T. F., Lauritsen, D., Cole, H. L., Beierle, K., and Chamberlain, E.: Performance of operational radiosonde humidity sensors in direct comparison with a chilled mirror dew-point hygrometer and its climate implication, *Geophys. Res. Lett.*, 30, 1860, doi:10.1029/2003GL016985, 2003.

Ware, R. H., Fulker, D. W., Stein, S. A., Anderson, D. N., Avery, S. K., Clark, R. D., Droege, K. K., Kuettner, J. P., Minster, J. B., and Sorooshian, S.: SuomiNet: A Real-Time National GPS Network for Atmospheric Research and Education, *B. Am. Meteorol. Soc.*, 81, 677–694, 2000.

Whiteman, D. N.: Examination of the traditional Raman lidar technique, I. Evaluating the temperature-dependent lidar equations, *Appl. Optics*, 42, 2571–2592, 2003a.

Whiteman, D. N.: Examination of the traditional Raman lidar technique, II. Evaluating the ratios for water vapor and aerosols, *Appl. Optics*, 42, 2593–2608, 2003b.

Whiteman, D. N., Melfi, S. H., and Ferrare, R. A.: Raman Lidar System for Measurement of Water Vapor and Aerosols in the Earth's Atmosphere, *Appl. Optics*, 31, 3068–3082, 1992.

Whiteman, D. N., Demoz, B., Di Girolamo, P., Comer, J., Veselovskii, I., Evans, K., Wang, Z., Cadirola, M., Rush, K., Sabatino, D., Schwemmer, G., Gentry, B., Melfi, S. H., Mielke, B., Venable, D., Van Hove, T., Browell, E., Ferrare, R., Ismail, S., and Wang, J.: Raman Water Vapor Lidar Measurements During the International H₂O Project. I. Instrumentation and Analysis Techniques, *J. Atmos. Ocean. Tech.*, 23, 157–169, 2006a.

Whiteman, D. N., Demoz, B., Di Girolamo, P., Comer, J., Veselovskii, I., Evans, K., Wang, Z., Cadirola, M., Rush, K., Sabatino, D., Schwemmer, G., Gentry, B., Melfi, S. H., Mielke, B., Venable, D., Van Hove, T., Browell, E., Ferrare, R., Ismail, S., and Wang, J.: Raman Water Vapor Lidar Measurements During the International H₂O Project, II. Instrument Comparisons and Case Studies, *J. Atmos. Ocean. Tech.*, 23, 170–183, 2006b.

Whiteman, D. N., Russo, Miloshevich, F. L., Demoz, B., Wang, Z., Veselovskii, I., Voemel, H., Hannon, S., Lesht, B., Schmidlin, F., Gambacorta, A., and Barnet, C.: Analysis of Raman lidar and radiosonde measurements from the AWEX-G field campaign and its relation to Aqua validation, *J. Geophys. Res.*, 111, D09S09, doi:10.1029/2005JD006429, 2006c.

Whiteman, D. N., Veselovskii, I., Cadirola, M., Rush, K., Comer, J., Potter, J., and Tola, R.: Demonstration Measurements of Water Vapor, Cirrus Clouds, and Carbon Dioxide Using a High-Performance Raman Lidar, *J. Atmos. Ocean. Tech.*, 24, 1377–1388, 2007.

Whiteman, D. N., Rush, K., Rabenhorst, S., Welch, W., Cadirola, M., McIntire, G., Russo, F., Adam, M., Venable, D., Connell, R., Veselovskii, I., Forno, R., Mielke, B., Stein, B., Leblanc, T., McDermid, S., and Vömel, H.: Airborne and Ground-based measurements using a High-Performance Raman Lidar, 27, 1781–1801, doi:10.1175/2010JTECHA1391, 2010.

Whiteman, D. N., Venable, D. V., and Landulfo, E.: “Comments on: Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements”, *Appl. Optics*, 1, 50, 2170–2176, 2011a.

Whiteman, D. N., Vermeesch, K. C., Oman, L. D., and Weatherhead, E. C.: The relative importance of random error and observation frequency in detecting trends in upper tropospheric water vapor, *J. Geophys. Res.*, 116, D21118, doi:10.1029/2011JD016610, 2011b.